

Corona-Sonderteil – Online-Lehre und Homeoffice am KIT  
*Corona Special – Online Teaching and Home Office at KIT*

Business Process Automation mit MS-Orchestrator  
*Business Process Automation with MS Orchestrator*

Peer Review 2.0  
*Peer Review 2.0*

## *Liebe Leserinnen und Leser,*

Computersimulationen und Datenanalysen beschleunigen Forschungsprozesse enorm und befördern den Erkenntnisgewinn. Wir freuen uns sehr, dass wir als eine der ersten Einrichtungen in Europa die neuartigen NVIDIA DGX A100 Systeme (Titelseite) zur Ermöglichung innovativer KI-Forschung in Betrieb nehmen konnten. Aber ohne Software würde die modernste Hardware nicht nutzbar sein. Und so ist es von großer Bedeutung, dass das Research Software Engineering (RSE) in Verbindung mit Open-Source-Software mehr und mehr ins Rampenlicht gerückt werden muss. Lesen Sie auf Seite 60, wie ein nachhaltiges Peer-Review-Verfahren auf Basis etablierter Prozesse in Open-Source-Softwareprojekten dazu beitragen kann.

Schnelllebigkeit und Agilität sind ständige Begleiter in einem modernen wissenschaftlichen Rechenzentrum. Es braucht mehr und mehr automatisierte Prozesse, um Plattformen und Infrastrukturen für Projekt- und Forschungsgruppen schnell und zuverlässig bereitzustellen. Auf Seite 28 erfahren Sie, wie das SCC durch Business Process Automation den Administrations- und Dokumentationsaufwand für die Erstellung und Abwicklung von Groupware-Ressourcen enorm reduziert hat.

Wie überall im Leben hatte die Corona-Pandemie auch Auswirkungen auf die Arbeitsabläufe am KIT. Dank der großen Anstrengung von Expertinnen und Experten aus dem SCC und vielen weiteren Organisationseinheiten des KIT konnten in kürzester Zeit neue, IT-gestützte Werkzeuge und – in die KIT-Infrastruktur integrierte – Dienste zur digitalen Zusammenarbeit in Forschung und zur Online-Lehre am KIT etabliert werden. Lesen Sie mehr dazu im ausführlichen Corona-Sonderteil (S. 4–12). Wir möchten uns an dieser Stelle für das Engagement und die erfolgreiche Zusammenarbeit aller Beteiligten bedanken!

Viel Freude beim Lesen.

*Martin Frank, Bernhard Neumair, Martin Nußbaumer, Achim Streit*

## *Dear reader,*

Computer simulation and data analysis accelerate research enormously and promote knowledge gain. We are very pleased that SCC is one of the first institutions in Europe to put the innovative NVIDIA DGX A100 systems (front page) into operation to enable innovative AI research. However, the most modern hardware would be worthless without software. Therefore it is ever more important to shine the light on research software engineering (RSE) and its connection with open source software. On page 60 you may read how a sustainable peer review workflow, based on established processes in open source software projects, can contribute to this.

Fast pace and agility are constant companions of a modern scientific data center. Increasingly automated processes are needed to provide platforms and infrastructures for projects and research groups and deliver them quickly and reliably. On page 28 you can read how SCC using Business Process Automation, greatly reduced the effort to administer and document the processes at SCC for allocation and handling of groupware resources.

Like in all our lives, the Corona pandemic also impacted life and work at KIT. Thanks to the concerted effort of SCC experts and many other organisational units of KIT, new IT-based tools and services for remote collaboration in research and online teaching were deployed and integrated with the KIT infrastructure in a very short time. Read more on this topic in the Corona special section (p. 4–12). We would like to take this opportunity to thank all contributors for their commitment and successful cooperation to secure the uninterrupted functioning of KIT during the Corona measures.

Enjoy reading.

*Martin Frank, Bernhard Neumair, Martin Nußbaumer, Achim Streit*

# Inhaltsverzeichnis

## CORONA-SONDERTEIL

- 04 Corona-Krise – Das KIT arbeitet im Homeoffice
- 05 MS Teams zur Kommunikation im Homeoffice
- 07 Ausbau der Netzinfrastruktur
- 09 Von der Präsenzuniversität zur temporären Online-Universität
- 12 Microsoft Teams in der Lehre
- 13 Service Support zum Semesterstart
- 13 Vorbereitung und Unterstützung der Lehrenden
- 14 Einführung von Zoom
- 16 Wissenwertes zum Opencast-Projekt im Corona-Sommersemester 2020
- 19 Mit ILIAS fit für die Online-Lehre
- 20 Online-Prüfungen in Zeiten der Pandemie

## DIENSTE UND INNOVATION

- 23 CAS Campus – zentrales Campus-Management-System am KIT
- 26 Akademische COMSOL Multiphysics Campus-Lizenz für das KIT
- 28 Business Process Automation mit MS-Orchestrator
- 31 Upgrade auf Microsoft Exchange 2019
- 33 Effiziente Vorgangsbearbeitung mit dem SCC-Ticketsystem
- 35 HoreKa!
- 36 bwUniCluster 2.0: Neues Tier-3 HPC-System am KIT
- 38 Anycast für die DNS-Resolver
- 40 Service Management in der European Open Science Cloud
- 44 bwSync&Share – der Landesdienst migriert zu Nextcloud
- 47 Digitalisierte Rechnungsverarbeitung am KIT

## FORSCHUNG UND PROJEKTE

- 49 Vorbereitung der Ginkgo-Softwarebibliothek für AMD-Grafikprozessoren
- 52 Die Helmholtz Metadata Collaboration Plattform
- 54 Wieso manchmal weniger mehr ist – auch in der Wissenschaft!
- 57 ‚Episteme in Bewegung‘ – Veränderlichem Wissen auf der Spur
- 60 Peer Review 2.0
- 63 Helmholtz Inkubator Pilotprojekt Uncertainty Quantification

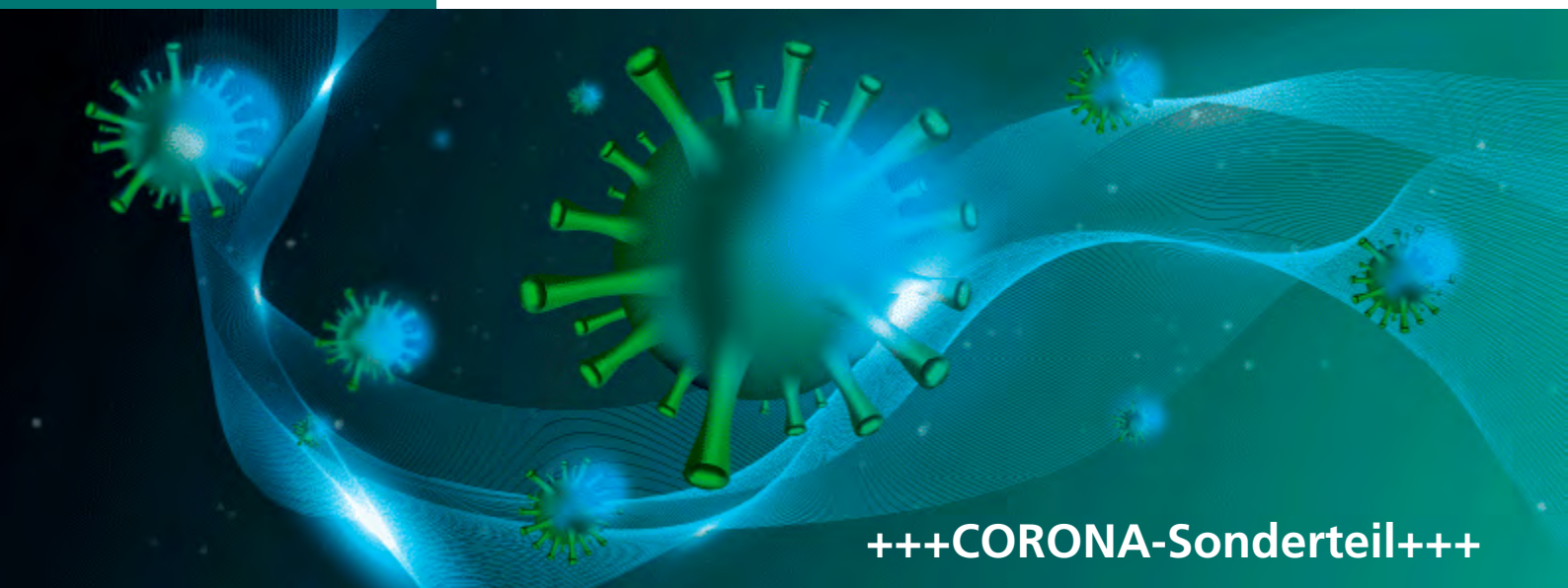
## STUDIUM UND WISSENSVERMITTLUNG

- 65 CAMMP goes online

## VERSCHIEDENES

- 67 Neues aus den SCC-Abteilungen
- 67 Impressum & Kontakte





## Corona-Krise – Das KIT arbeitet im Homeoffice

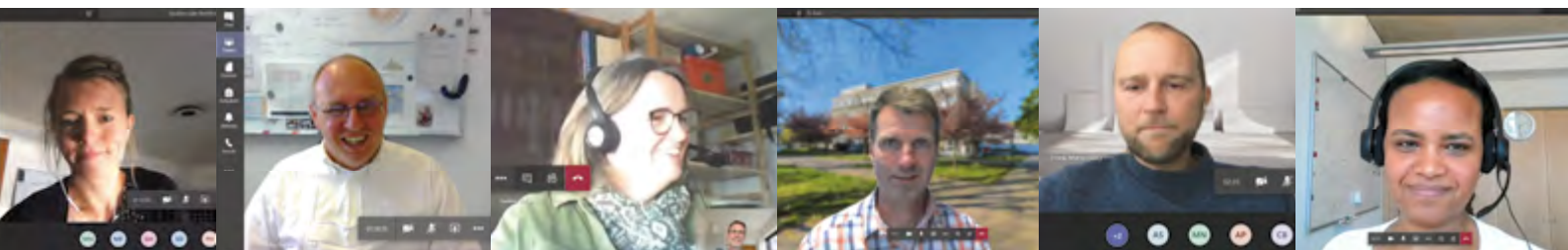
*Achim Grindler*

Aufgrund der Corona-Verordnungen des Landes Baden-Württemberg hat auch das KIT umfangreiche Maßnahmen ergriffen, um die Ausbreitung der Epidemie einzudämmen und dennoch weitestgehend arbeitsfähig zu bleiben. Ein wesentlicher Baustein bei der Eindämmung der Virusausbreitung ist die gegenseitige Distanzwahrung von Beschäftigten und Studierenden. Das SCC hat sich daher ab Mitte März, noch in der vorlesungsfreien Zeit, auf den schnell wachsenden Bedarf von Homeoffice-Arbeitsplätzen mit den technischen Möglichkeiten für eine Zusammenarbeit vorbereitet. Diesen Bedarf hatte das SCC zum einen selbst, zum anderen war es aber auch ein gefragter Partner für die Einführung neuer digitaler Werkzeuge und Anpassung von Arbeitsabläufen am KIT in Zeiten von Corona.

Da ein großer Teil der Beschäftigten nun von Zuhause aus arbeitete, hat das SCC sowohl die bestehenden Angebote für den Remote Zugriff auf KIT-Ressourcen ausgebaut als auch Teile der Homeoffice-Arbeitsplätze mit Geräten ausgestattet. Dienste, die zur Verwaltung der Arbeitsplatzrechner am KIT im Einsatz waren, wurden für Geräte im Heimnetz, also über das Internet, nutzbar gemacht. Eine in die IT-Infrastruktur des KIT integrierte Plattform für die Online-Zusammenarbeit mit Telefonie-, Videokonferenz- und Chat-Funktion war am KIT nicht flächendeckend ausgerollt. Eine solche Lösung (Seite 5) wurde zeitnah evaluiert sowie deren Produktivsetzung geklärt. Nicht zuletzt galt es, Hinweise zum Arbeiten im Homeoffice zu kommunizieren und dafür erweiterte Supportstrukturen aufzubauen. Dienste, die eine Präsenz von

Personen an den Standorten erforderten, wurden kurzfristig geeignet umgestellt oder ausgesetzt. Dies betraf u.a. die Präsenzberatung, die Druckausgabe und die Identitätsprüfung von Personen bei der Beantragung von digitalen Zertifikaten.

Parallel zu diesen Aktionen bildeten sich Arbeitsgruppen, um das KIT auf den bevorstehenden Online-Vorlesungsbetrieb im Sommersemester ab 20. April vorzubereiten. Neben dem stabilen Betrieb des ILIAS E-Learningsystems und der OpenCast Streamingplattform standen Videokonferenzsysteme für Seminare, Vorlesungen und Prüfungen im Fokus. Mehr dazu ab Seite 9.



## MS Teams zur Kommunikation im Homeoffice

Das KIT hat die Entwicklung der Corona-Situation frühzeitig beobachtet und schrittweise Einschränkungen zum Umgang mit dieser neuen Realität eingeführt. Am 16. März erging schließlich die Entscheidung, „die Möglichkeiten zu mobilem Arbeiten von zuhause auszuschöpfen“.

*Patrick von der Hagen, Ulrich Weiß, Jörg Kramer*

Diese Entwicklung war absehbar und so hat sich das SCC schon im Vorfeld intensiv mit der Frage beschäftigt, wie kurzfristig die dann aus dem Homeoffice arbeitenden Mitarbeiterinnen und Mitarbeiter bei der Zusammenarbeit wirksam unterstützt werden könnten.

Angesichts der Erfahrungen mit mobiler Arbeit und mit Telearbeit war klar, dass eine kurzfristige Unterstützung mit den vorhandenen Strukturen und Ressourcen nicht möglich sein würde. Einige Dienste, wie beispielsweise die Webauftritte, waren durch die neue Situation nicht wesentlich betroffen, andere mussten mit höheren Anwenderzahlen kämpfen. Insgesamt entstanden für das Homeoffice-Szenario jedoch völlig neue Anforderungen und Randbedingungen an das Dienste-Portfolio des SCC.

Diese stellten sich wie folgt dar:

- Eine Telefonie-Lösung für das Homeoffice war kurzfristig nicht für alle KIT-Angehörigen verfügbar
- Bei der Durchführung von Videokonferenzen über den Dienst „DFNconf“ des DFN wurden frühzeitig Kapazitätsengpässe festgestellt
- Der Chat-Dienst des SCC auf Basis von XMPP hat sich in der Breite bislang nicht etabliert
- Eine integrierte Lösung, um z. B. im Chat spontan eine Sprachverbindung aufzubauen und gemeinsam ein Dokument zu betrachten, war ebenfalls nicht verfügbar

Eine SCC-Arbeitsgruppe beschäftigte sich bereits – im Rahmen laufender Vertragsverhandlungen zu den Microsoft Clouddiensten – mit der Integration von Clouddiensten in das Dienste-Portfolio

des SCC. Eine in diesem Zusammenhang entstandene Teststellung konnte so mit verstärktem Einsatz schnell zur Bewertung auf Eignung hinsichtlich der sich abzeichnenden Corona-Beschränkungen untersucht werden. Dabei wurde insbesondere eine prototypische Anbindung des Verzeichnisdienstes Azure-AD an das am SCC verwendete Produkt ForgeRock Identity Management (IDM) sowie an das Shibboleth Single-Sign-On-System (SSO) realisiert. Aufgrund der so gesammelten Erfahrungen konnte das SCC eine kurzfristige Integration der Kollaborationsplattform Microsoft Teams in die IT-Infrastruktur des KIT in Aussicht stellen. Die Nutzung von MS Teams versprach nicht nur hinsichtlich ökonomischer Aspekte<sup>1</sup> Vorteile, sondern war auch sofort verfügbar und bedurfte keiner zusätzlichen Beschaffung. Vor dem Hintergrund der rasanten Verschärfung der Corona-Krise bildete diese Zeiteinsparung den entscheidenden Faktor, der die Beschäftigten des KIT unmittelbar arbeitsfähig hielt und die notwendige Arbeitsplattform zur Vorbereitung des digitalen Online-Semesters für 4.000 Dozierende und 25.000 Studierende darstellte.

Bereits am 12. März fiel im SCC die Entscheidung, dem Corona-Notfall-Stab des KIT MS Teams als Kollaborationsplattform vorzustellen und die kurzfristige Einführung der Plattform vorzuschlagen. Der Nutzerkreis der Testumgebung wurde ausgehend von den Krisenstab- und Taskforce-Mitgliedern sukzessive erweitert, um einen kontinuierlichen Abgleich notwendiger Anforderungen mit der fortschreitenden Entwicklung zu erreichen. Parallel dazu wurden Gespräche hinsichtlich Datenschutz und Informationssicherheit geführt sowie die Weiterentwicklung der Anbindung forciert, um

neben der prototypischen Testinstanz auch eine Anbindung der produktiven Umgebung zu ermöglichen. Weiterhin musste ein Support-Team aufgebaut, Dokumentation vorbereitet, Software-Pakete für die zentrale Softwareverteilung mittels OPSI vorbereitet sowie Policy- und Design-Entscheidungen für die Azure-AD-Umgebung getroffen werden.

In Absprache mit der Datenschutzbeauftragten entschied man sich für eine Duldung der Cloud-Nutzung während der Corona-Krise, wenn die Nutzen den umfassend informiert werden und persönlich zustimmen. Das SCC hat dazu das Self-Service-Portal [my.scc.kit.edu](https://my.scc.kit.edu) um die nötigen Funktionen erweitert. Durch die eingeschränkte Duldung war es nötig, den vielfältigen Nutzungsumfang der Microsoft Clouddienste auf MS Teams zu beschränken.

Nach diesen umfangreichen Vorarbeiten konnte die Anbindung von MS Teams für Mitarbeitende schließlich am Abend des 19. März, und damit drei Tage nach dem Lockdown, durchgeführt werden. Der IT-Expertenkreis des KIT wurde am 20. März über die Einführung informiert. Das nächste Treffen des Gremiums fand dann schon am 25. März mit über 200 Teilnehmenden erfolgreich über MS Teams statt. Trotz eingeschränkter Unterstützungsmöglichkeiten im Homeoffice nutzten bereits eine Woche nach der Einführung 25% der Mitarbeitenden die neue Plattform.

Als nächster Meilenstein wurde die Unterstützung der Gäste und Partner des KIT vorbereitet und am 26. März durchgeführt.

<sup>1</sup> MS Teams ist Teil des Landesvertrags „Microsoft Enrollment for Education Solutions“, dem das KIT beigetreten ist.

In der folgenden Woche stand die Umstellung der Studierenden an. Bereits in der Vergangenheit konnten Studierende in einem eigenen Mandanten Cloud-dienste von Microsoft beispielsweise im Zusammenhang mit den Azure Dev Tools for Teaching (ADT4T) nutzen. Dieser von den Mitarbeitenden getrennte Mandant wurde beibehalten und ebenfalls an das Identitätsmanagement des KIT angebunden. Leider lässt Microsoft die Anbindung von zwei Mandanten an denselben Shibboleth-Dienst nicht zu, so dass in diesem Fall noch ein technischer Umweg unter Nutzung des „Keycloak“-Authentisierungsservers eingerichtet werden musste. Da dieser Dienst den Shibboleth-IDP des KIT zur Authentifizierung einbindet, ist ein Single-Sign-On für die Studierenden weiterhin gewährleistet. Die Umstellung der Studierenden konnte schließlich am 3. April durchgeführt werden.

Durch die Trennung in die beschriebenen zwei Mandanten ist für die Studierenden eine Anmeldung am Mitarbeiter-Mandanten nicht ohne weiteres möglich. Dazu ist zunächst eine vom Mitarbeiter-Mandanten initiierte Einladung als Gast erforderlich, die später für alle Studierenden automatisiert wurde. Dies ist im Artikel zum Einsatz von MS Teams in der Lehre weiter ausgeführt.

Insgesamt wurden am KIT innerhalb von drei Wochen

- die Nutzung von MS Teams mit der Datenschutzbeauftragten geklärt
- ein Self-Service für die nötige Zustimmung bereitgestellt
- ein Support-Team etabliert und die notwendigen Anleitungen erstellt
- die Anbindung für alle Mitglieder des KIT über IDM und SSO entwickelt
- bestehende Konten korreliert und verknüpft
- MS Teams als Plattform erfolgreich eingeführt

Insgesamt ist MS Teams als Plattform trotz des engen Zeitplans und eingeschränkter Informations- und Schulungsmöglichkeiten sehr gut aufgenommen worden und hat sich schnell bewährt. Dabei reicht die Nutzung von der Projektarbeit bis hin zum täglichen Plausch an der virtuellen Kaffeemaschine. Viele Anwender wünschen auch eine zukünftige Nutzung dieses inzwischen etablierten Tools. Daher werden inzwischen erste Gespräche zu einer langfristigen Nutzung und einer tiefergehenden Integration in die Dienste des KIT geführt. Denkbar wäre eine Erweiterung um Telefoniefunktionen durch eine Anbindung an die vorhandene Telefonanlage.

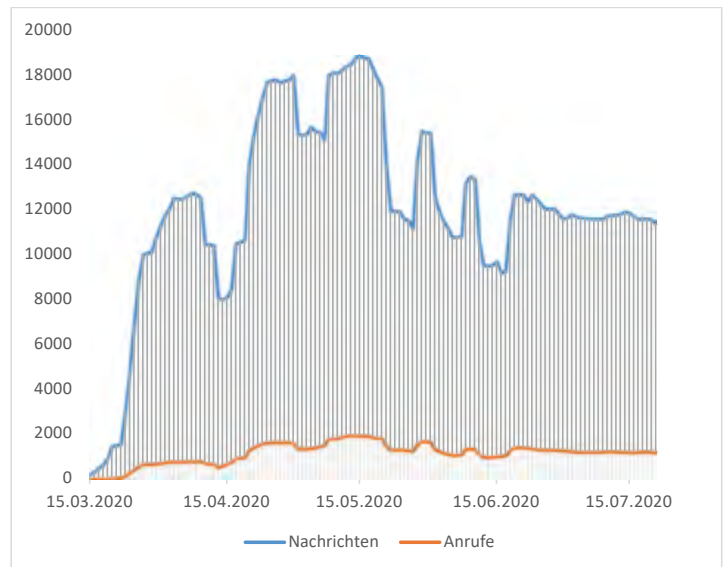


Abbildung 1: Stabilisierung der Nutzung auf hohem Niveau, trotz Wiederaufnahme des regulären Dienstbetriebs

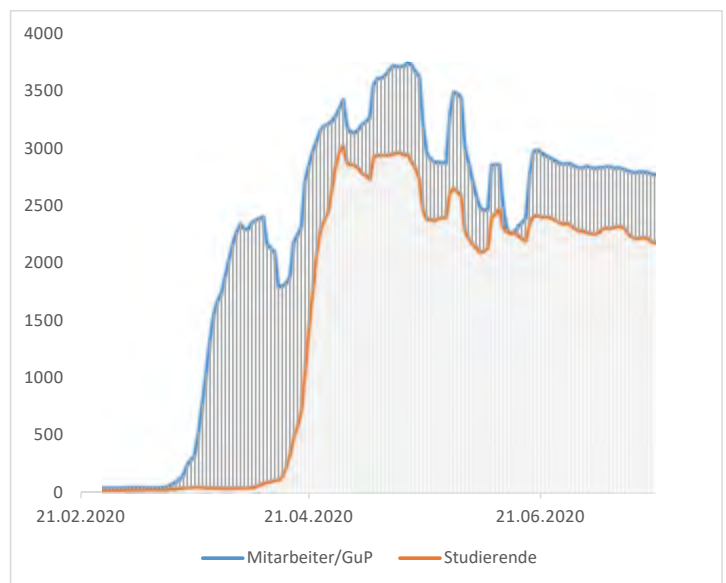


Abbildung 2: Anzahl der täglichen Nutzer (Mitarbeiter-Mandant)

### Einschränkungen von MS Teams am KIT

Die Duldung durch den Datenschutz beschränkt sich auf MS Teams als Kommunikationsplattform im engeren Sinn. Daraus ergeben sich in der Praxis einige Einschränkungen. Beispielsweise ist es ohne SharePoint-Lizenz nicht möglich, im Chat Dateien mit den Kollegen zu teilen und ohne Exchange-Integration fehlt eine Kalenderfunktion zur Besprechungsplanung.

## Ausbau der Netzinfrastruktur

Bedingt durch die Corona-Krise gab es einige Herausforderungen, die sich aus Szenarien wie Homeoffice und Online-Lehre ergaben. Hierauf war das Netzwerk des KIT gut vorbereitet. Der VPN-Dienst konnte in kürzester Zeit massiv aufgerüstet werden. Die Kapazitäten im Netzwerk waren durch die vorausschauende Planung in der Vergangenheit bereits ausreichend oder konnten kurzfristig erhöht werden.

*Klara Mall, Benedikt Neuffer, Julian Schuh*

Mit dem Beginn der Corona-Krise ergaben sich weitreichende Herausforderungen für die Netzwerkinfrastruktur des KIT. Die KIT-Mitarbeiter arbeiteten vermehrt aus dem Homeoffice, darüber hinaus sollte für die Studierenden das Sommersemester komplett digital stattfinden, bei den Mitarbeitern stand das Thema VPN – für den Remote-Zugriff auf die Ressourcen am KIT – sowie Video Conferencing im Vordergrund. Bei den Studierenden und Lehrenden das Zurverfügungstellen bzw. Abrufen von Online-Lehrinhalten sowie Live-Vorträge und mündliche Prüfungen.

### Aufrüstung des VPN-Dienstes

Am schnellsten war die Aufrüstung des VPN-Dienstes erforderlich, da der Wechsel ins Homeoffice plötzlich kam und nicht wie das Online-Studium bis zum Vorlesungsbeginn vorbereitet werden konnte. Der OpenVPN-Dienst wird für

den sicheren Zugriff auf das KIT-Netz von extern aus genutzt und in Form mehrerer virtueller Maschinen hinter dem Load Balancer des SCC betrieben. Bisher waren für den Standard-VPN-Dienst vier virtuelle Maschinen (VMs) im Einsatz, wobei je Maschine 253 IPv4-Adressen für VPN-Clients zur Verfügung gestellt wurden. In der Spitze wurden bis zu diesem Zeitpunkt maximal 500 gleichzeitige VPN-Nutzer gemessen. Im Kontext der aktuellen Lage musste jedoch von mehr als 1000 gleichzeitigen Nutzern des Dienstes ausgegangen werden, wofür schon alleine die vorhandenen IPv4-Adressen nicht ausgereicht hätten. Zukünftig sollen alle Systeme und Nutzer des KIT via IPv6 angebunden werden, wodurch auf die Vergabe von IPv4-Adressen verzichtet werden kann.

Die Hochrüstung geschah in enger Zusammenarbeit mit dem VM-Team am Wochenende, bevor eine große Anzahl

der Beschäftigten ihre Arbeit aus dem Homeoffice begannen. Dies war leider nicht sehr weit im Voraus abzusehen, so dass hier sehr kurzfristig reagiert werden musste. Es wurden sechs weitere OpenVPN-Maschinen in Betrieb genommen, wobei jede Maschine 1021 IPv4-Adressen für VPN-Clients bereitstellte. Bereits am folgenden Montag wurden über 1000 gleichzeitige Nutzer im VPN gemessen, so dass die kurzfristige Reaktion als Erfolg angesehen werden kann. In der Hochphase der Arbeit von zu Hause aus hat sich die Zahl der gleichzeitigen VPN-Nutzer bei ca. 2500 in der Spitze eingependelt, der Gesamt-Durchsatz auf etwa 1Gbit/s. Die CPU-Last auf den einzelnen VPN-Maschinen konnte durch die kurzfristigen Skalierungsmaßnahmen im grünen Bereich gehalten werden.

Die schnelle und erfolgreiche Hochrüstung wurde durch den Einsatz der Open-Source Software OpenVPN (keine Lizenzierung erforderlich) und des Open-Source Automatisierungs-Werkzeugs Ansible sowie der exzellenten Zusammenarbeit mit dem VM-Team ermöglicht.

Die genannten Zahlen beziehen sich auf den Standard-VPN-Zugang. Auch der VPN2VLAN-Dienst, der den sicheren Zugang in Institutsnetze ermöglicht, wurde aufgerüstet. Aus technischen Gründen war dies leider nur in begrenztem Maße möglich, wobei weitere Verbesserungen an dieser Stelle bereits geplant sind. Unter normalen Umständen wurde der VPN2VLAN-VPN-Dienst im Schnitt von ca. 200 gleichzeitigen VPN-Nutzern verwendet, durch die Corona-Maßnahmen stieg diese Zahl auf ca. 1000. Hierbei ergab sich die besondere Situation, dass sowohl

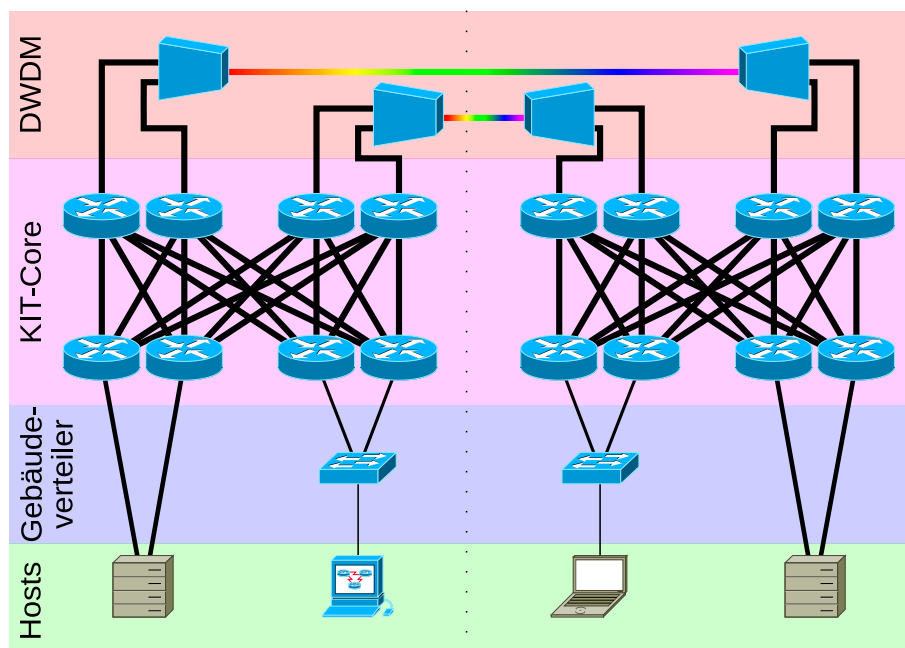


Abbildung 1: Neuer KIT-Core





Abbildung 2: Kürzlich abgeschalteter Router neben einem noch laufenden Router des alten KIT-Backbones.

neue VPN2VLAN-Zugänge eingerichtet als auch bestehende Zugänge mit weiteren IPv4-Adressen erweitert werden mussten, was in einigen Fällen dazu führte, dass völlig neue Subnetze eingerichtet werden mussten.

Insgesamt wurde in dieser Zeit ein starker Anstieg des Support-Aufwands für den VPN-Dienst registriert, da viele den VPN-Dienst erstmals nutzten und darüber hinaus, wie bereits beschrieben, neue Zugänge eingerichtet und vorhandene erweitert werden mussten. Dies konnte aber durch den hohen Einsatz der Support-Teams gut bewältigt werden.

### Kapazität des KIT-Netzwerks

Für die oben genannten Anwendungen wie Video Conferencing und Online-Lehre muss genügend Kapazität im Netzwerk bereit stehen. Im Homeoffice-Szenario findet die Kommunikation in der Regel vom Internet in das KIT und umgekehrt statt. Insgesamt waren wir hierbei sehr gut aufgestellt: Die Außenanbindung des KIT hat schon seit längerer Zeit eine Bandbreite von insgesamt 400 Gigabit/s. Wie in dem Artikel „Neuer Backbone“ (SCC-News 1/2019) beschrieben, befindet sich derzeit der neue 100G-KIT-Core im Aufbau, wobei hier inzwischen gute Fortschritte erzielt werden konnten. Durch großen Einsatz der Netzabteilung und sehr engen Austausch mit dem Herstel-

ler konnte dieser ein Software-Problem, das in der Vergangenheit wiederholt für größere Störungen sorgte, rechtzeitig vor Corona beheben. Die Hosts der VMware-Virtualisierungsumgebungen sind schon seit längerer Zeit mit jeweils 2x25G an den KIT-Core angeschlossen.

Während das Netzwerk durch den neuen KIT-Core bandbreitentechnisch bereits sehr gut aufgestellt ist, konnten wir bei weiteren Migrationen punktuell genau die Stellen priorisieren, die nun im Rahmen der Corona-Krise für die oben genannten Anwendungen von zentraler Bedeutung sind. Hierbei handelt es sich konkret um die Anbindung der Opencast-Infrastruktur (Seite 16) mit 40G an den neuen KIT-Core. Gleichzeitig wurde die Anbindung des KIT-Core an die Border-Router (Internetanbindung) auf 100G hochgerüstet. Weiterhin wurden die Load Balancer mit 40G an den KIT-Core migriert und der Fileserver am Campus Nord an den KIT-Core umgezogen, so dass auch hier eine höhere Bandbreite zur Verfügung steht. Die einzige verbleibende Engstelle hätte die in die Jahre gekommene zentrale Firewall darstellen können, welche lediglich einen Durchsatz von 10 Gbit/s ermöglicht. Für Anwendungen mit hohen Bandbreitenanforderungen bei der Kommunikation mit Systemen im Internet wurde daher eine Lösung implementiert, welche den Datenverkehr unter Umgehung der Firewall ins Internet leitet – so

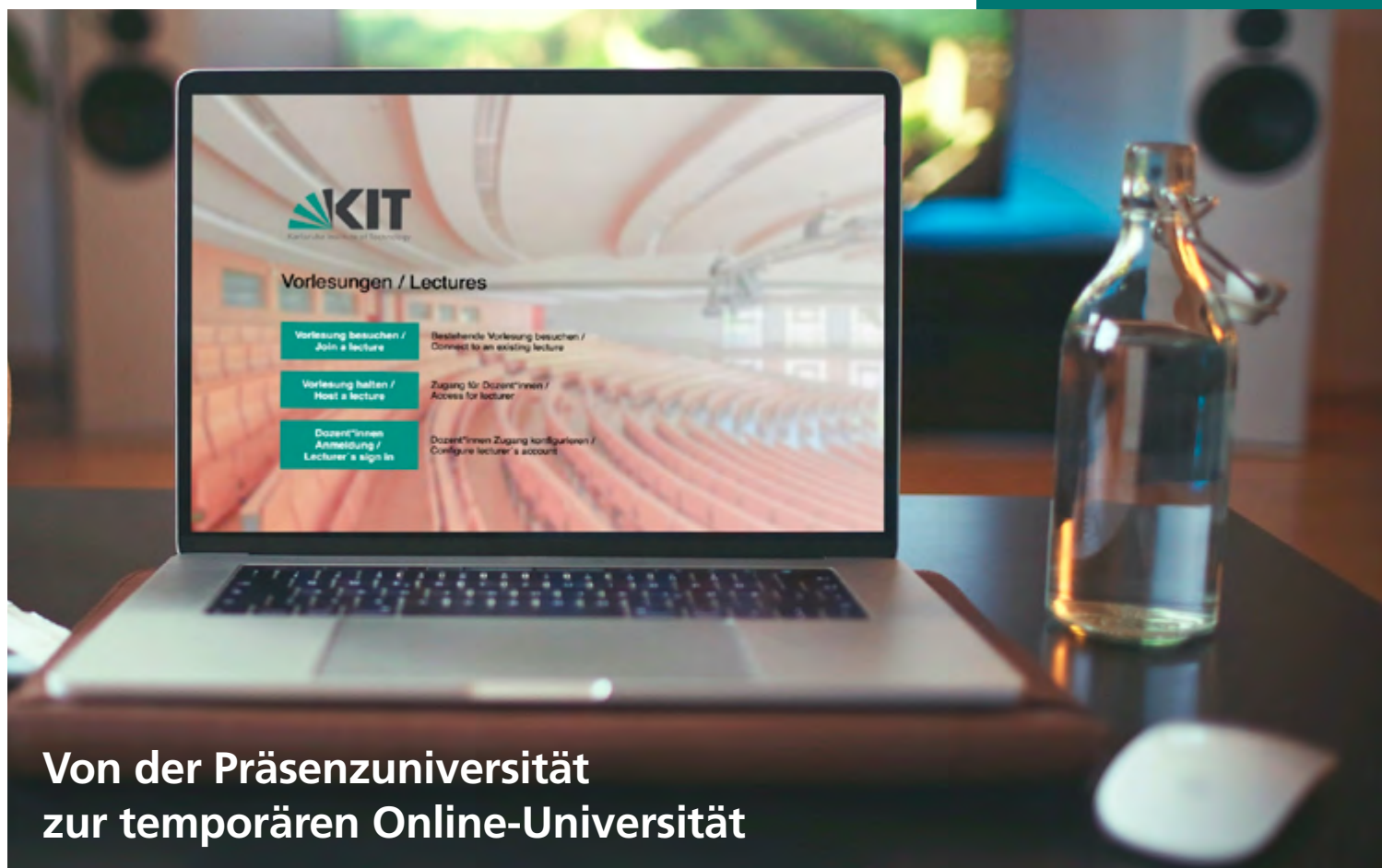
beispielsweise bei der oben genannten Opencast-Infrastruktur.

Mit diesen Optimierungen, die im Rahmen der Migration ohnehin geplant waren, war das Netzwerk des KIT sehr gut für die genannten Online-Szenarien aufgestellt. Es konnten keine Engpässe festgestellt werden; das Netzwerk war den stark gestiegenen Anforderungen problemlos gewachsen. Generell konnten wir trotz Corona weitere Fortschritte bei der Migration auf den neuen KIT-Core erzielen. Das letzte fehlende Stück der Hochrüstung des Netzwerks – die zentrale Firewall – soll noch in diesem Jahr beschafft werden. Diese wird einen Durchsatz von 2x60G ermöglichen.

### Enhancement of the KIT Network

The KIT network was well prepared for the challenges of the Corona crisis. The VPN service has been massively upgraded in a very short time. The network capacity was already sufficient due to proactive planning in recent years, or could be upgraded easily. This article describes in detail the challenges we faced and the changes that have been made in the network infrastructure.





## Von der Präsenzuniversität zur temporären Online-Universität

Durch die Corona-Krise war das KIT herausgefordert, die Lehre innerhalb von fünf Wochen auf eine reine Online-Lehre umzustellen. Wie diese Mammutaufgabe gelungen ist und was dafür in kürzester Zeit realisiert wurde, zeigt dieser Artikel. Der Fokus liegt dabei auf den zentralen Maßnahmen, die nur wirksam waren, da von den Fakultäten und den Lehrenden dezentral gleichermaßen Großartiges geleistet und die geschaffenen Möglichkeiten bereitwillig in Gebrauch genommen wurden. Stimmen in den Sozialen Medien, Rückmeldungen im Semester und erste Rückläufer aus der Lehrevaluation deuten auf einen gelungenen Lehrbetrieb unter Corona-Bedingungen hin.

*Andreas Sexauer<sup>1</sup>, Martin Nussbaumer*

Am KIT, als Präsenzuniversität, wurde die Lehre bisher durch digitale Angebote angereichert, um beispielsweise die Qualität zu steigern, geographische Distanzen zu überwinden oder die Chancengleichheit zu verbessern. In der ersten Märzhälfte wurde durch den Corona-Erlass des Landes Baden-Württemberg der Vorlesungsbetrieb an allen Hochschulen bis 20. April ausgesetzt. Damit war schnell klar, dass zu diesem Zeitpunkt der Start des regulären Vorlesungsbetriebs nicht garantiert werden kann. Der Vizepräsident für Lehre und akademische Angelegenheiten, Alexander Wanner, hat da-

raufhin den IV-Bevollmächtigten Hannes Hartenstein beauftragt, eine Experten-Gruppe Online-Lehre zu initiieren. Die aus Mitgliedern des Gremiums MIK (Medien- und IV-Infrastruktur-Koordinierung)<sup>2</sup> bestehende Gruppe sollte prüfen, ob und wie ein Online-Lehrbetrieb aufgenommen werden kann. Die Expertengruppe traf sich am 13. März erstmalig vor Ort, alle weiteren Treffen fanden rein online statt. Auch hier war gefordert, die Arbeit aus dem Homeoffice heraus zu realisieren und sich entsprechend zu organisieren. Sarah Holstein und Andreas Sexauer vom Zentrum für Mediales Lernen (ZML)

wurden als Durchführungsverantwortliche bestimmt. Die Gruppe erkannte, dass das Vorhaben nur dann skalieren kann, wenn die didaktischen Szenarien eng mit der technischen Umsetzung abgestimmt und bereitgestellt werden. Abschätzungen für die Realisierung der notwendigen IT-Infrastrukturen und Werkzeuge ergaben, dass ohne eine Diversifizierung mit On-Premise- und Cloud-Strukturen kein stabiler Betrieb zu gewährleisten ist. Frust bei allen Beteiligten, weil Systeme unter der Last zusammenbrechen und nicht zuverlässig funktionieren, sollte auf jeden Fall vermieden werden. Schließlich galt es,

<sup>1</sup> Zentrum für Mediales Lernen (ZML)

<sup>2</sup> Mitglieder der MIK sind KIT-Bibliothek, Zentrum für Mediales Lernen und das Steinbuch Centre for Computing

die Szenarien und zugehörigen Werkzeuge plötzlich nicht nur in einzelnen Lehrveranstaltungen, sondern flächendeckend mit über 24.000 Studierenden und mehr als 5.000 Personen in Lehre und Forschung zu nutzen. Gleichzeitig sollte in der besonderen Situation eine klare

Kommunikation und Strukturierung erfolgen, um der pandemiebedingt ohnehin hohen Änderungsdynamik zu begegnen. Am KIT gab es bereits mit dem Lernmanagement-System ILIAS eine solide Basis, um Lehrangebote online zu begleiten. Weitere Vorerfahrungen gab es z.B. durch

das Vorlesungsaufzeichnungsprojekt, DFN-Videokonferenzen oder erste Tests mit der Integration von Microsoft-Cloud-Produkten. Für die Videodistribution und den verstärkten Einsatz von Videokonferenzen galt es, in kürzester Zeit Systeme zu beschaffen und in einen flächende-

Szenario	S1 Begleitetes Selbststudium mit ILIAS	S2 Interaktive Lehrveranstaltungen / kleine Gruppen	S3a Große Lehrveranstaltungen ohne Live-Interaktion (asynchron)	S3b Große Lehrveranstaltungen mit Live-Interaktion (synchron)
	Lehrende stellen Materialien wie z.B. kommentierte Foliensätze und Skripte in ILIAS zur Verfügung. Die Interaktion erfolgt über Foren und E-Mail.	Lehrende halten Seminare und ähnliche Lehrveranstaltungen über Microsoft Teams oder Zoom live von zu Hause oder aus dem Büro ab. Studierende können per Text-Chat oder Videokonferenz interagieren.	Lehrende zeichnen mit PowerPoint oder per Bildschirmaufzeichnung die Vorlesung vorab am eigenen Computer auf und stellen diese mit Opencast direkt im ILIAS-Kurs ihren Studierenden zur Verfügung. Die Interaktion erfolgt über Foren oder E-Mail.	Lehrende halten die Vorlesung live mit Zoom von zu Hause oder aus dem Büro ab. Die Vorlesung kann optional aufgezeichnet und zusätzlich wie in S3a bereitgestellt werden. Studierende können im Zoom-Raum per Text-Chat interagieren.
<b>Verwendete Systeme</b>	ILIAS	ILIAS Microsoft Teams Zoom	ILIAS PowerPoint/OBS Opencast	ILIAS Zoom (optional Opencast)

Tabelle 1: Szenarien für die Online-Lehre im Sommersemester

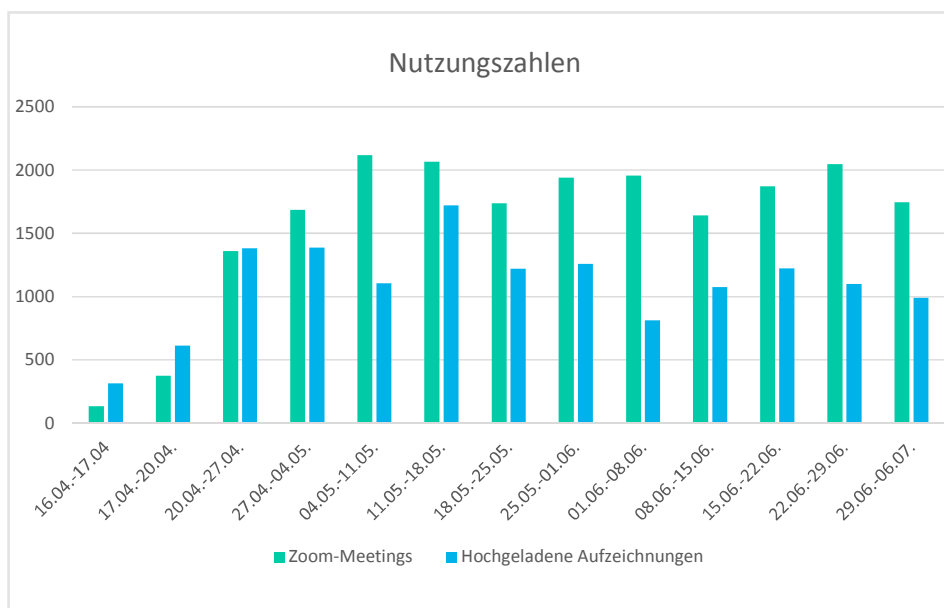


Abbildung 1: Nutzungszahlen ausgewählter Systeme für die Szenarien große Lehrveranstaltung ohne und mit Live-Interaktion

ckenden Betrieb zu überführen. Dabei mussten die Aktivitäten zum einen so koordiniert werden, dass die Lehrenden möglichst früh an die Werkzeuge herangeführt werden konnten, um das Semester vorzubereiten, und zum anderen gleichzeitig Zeiträume für die Implementierung der Systeme genutzt werden konnten. All dies erfolgte in kurzen Etappen jeweils auf Basis von Risikoabschätzungen. Die Expertengruppe definierte wenige Szenarien (s. Tabelle 1), mit denen ein großer Teil der Lehre abgedeckt werden konnte, und evaluierte die dafür passende Technik. Alle diese Szenarien teilen den Einsatz von ILIAS als führendes System für die Organisation, damit Studierende möglichst alles an einem Ort vorfinden.

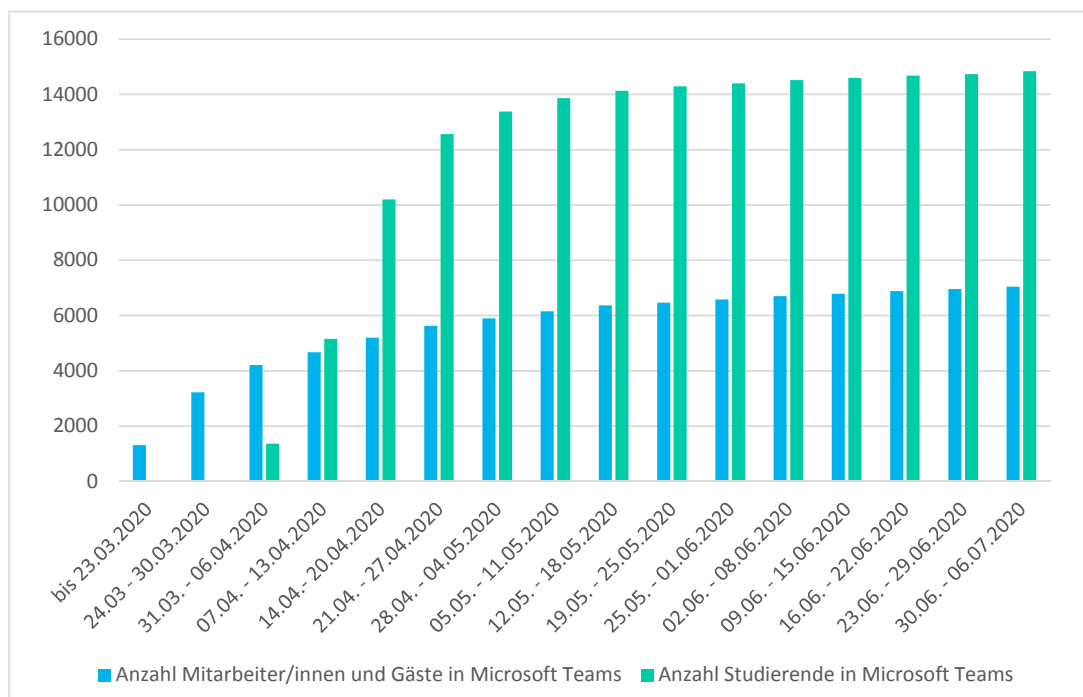


Abbildung 2: Freischaltung der Zugriffsmöglichkeit auf Microsoft Teams über die Zustimmung zur Cloud-Nutzung

Leitgedanken für alle Szenarien der Expertengruppe waren:

- Einfachheit und Klarheit
- Sicherstellung von Skalierbarkeit und Zuverlässigkeit
- Wo möglich Aufzeichnungen statt Live-Veranstaltungen
- Werben für Geduld und eine umsichtige Nutzung

Insgesamt konnte mit den Maßnahmen am KIT gut in das Sommersemester gestartet werden. Es entstand eine positive Grundstimmung und ein Gefühl, sich der

Krise gemeinsam zu stellen. Die bereitgestellten Szenarien bewährten sich in der Praxis, und die Infrastruktur war dank der vorherigen Kapazitätserweiterungen von ILIAS und der Ergänzung durch Cloud-Systeme stabil und performant. Durch die intensivere Nutzung der IT-Infrastruktur gab es zwar mehr Supportanfragen als gewöhnlich, diese konnten aber nicht auf generelle Probleme zurückgeführt werden. Die momentan zu großen Teilen online stattfindende Lehre soll und kann die präsenste Lehre nicht ersetzen, stellt aber für den Zeitraum der Krise eine solide Alternative dar. Erste Rück-

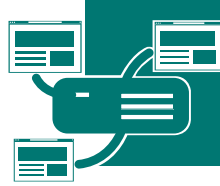
meldungen aus der Lehrevaluation und von Studierenden bestätigen dies. Die Nutzungszahlen der Systeme deuten auch darauf hin, dass die Szenarien über das Semester kontinuierlich genutzt wurden. So gibt es beispielsweise pro Woche ca. 1600 Zoom-Meetings (S. 14) und etwa 1100 Uploads von Aufzeichnungen für Lehrveranstaltungen in ILIAS (S. 19). Die entwickelten Szenarien stellen eine gute Voraussetzung für das kommende Wintersemester dar, bei dem damit zu rechnen ist, dass es wiederum in großen Teilen digital stattfinden wird.

## Das Online-Semester in Zahlen

Im Sommersemester gab es **3.634** Lehrveranstaltungen in Campus-Portal. Es wurden **2.378** Kurse im KIT-Lernmanagementsystem ILIAS und **742** Microsoft Teams für die Lehre eingerichtet. In die Nutzung von Microsoft Teams haben **7.195** Mitarbeiter/innen und **14.983** Studierende eingewilligt. In Opencast wurden **16.770** Aufzeichnungen für die Verwendung in den ILIAS-Kursen hochgeladen. Über Zoom fanden **24.882.681** Meeting-Minuten statt. In der KIT-Umgebung für mündliche Prüfungen wurden **1679** Zeit-Slots gebucht.

Zahlen: Expertengruppe Online-Lehre, Stand 20.07.2020





## Microsoft Teams in der Lehre

Die Cloud-Plattform Microsoft Teams wurde vorgesehen, um das Szenario S2 „Interaktive Lehrveranstaltungen in kleinen Gruppen“ für die Online-Lehre zu unterstützen. Ein über diese Plattform eingerichtetes „Team“ ist eine dauerhaft verfügbare Arbeitsumgebung, die Chat, Diskussion in Foren und auch Videokonferenzen in Form von Meetings anbietet. Im Gegensatz zum Szenario „Große Lehrveranstaltungen“ sind die Einschränkungen der Teilnehmerzahlen und Moderationsmöglichkeiten bei Teams von geringerer Bedeutung. Im Vordergrund steht die Kollaboration in der Gruppe, wie sie z.B. in Seminaren stattfindet, die von MS Teams besonders gut unterstützt wird.

Michael Meier, Andreas Sexauer<sup>1</sup>

Da die technische Anbindung von MS Teams sowie Datenschutzaspekte bereits im Zusammenhang mit dem Einsatz von MS Teams im Homeoffice betrachtet wurden (S. 5), konzentrierten sich die Herausforderungen beim Einsatz in der Lehre auf organisatorische Regeln, die Erstellung und Unterstützung begleitender Prozesse sowie den didaktisch sinnvollen Einsatz der neuen Werkzeuge in der Lehre.

Als zentrale Anlaufstelle für die Verwaltung aller Lehrveranstaltungen dient das Campus Management. Hier wurde kurzfristig für die Dozierenden die Möglichkeit geschaffen, veranstaltungsbegleitende Teams einzurichten und zu verknüpfen. Inzwischen sind auf diesem Weg über 700 Teams für die Lehre eingerichtet worden (Abbildung 1).

Dabei werden automatisch Einladungs-Links erzeugt, die die Dozierenden an die Studierenden kommunizieren können. Sofern die Studierenden der Nutzung von MS Teams zugestimmt haben, ist über diesen Weg ein einfacher Beitritt möglich (Abbildung 2). Dies war notwendig, da am KIT Dozierende und Studierende in getrennten Teams-Umgebungen – sogenannten Tenants – organisiert sind. Für den Einsatz in der Lehre mussten Studierende einen Zugang als Gast im Mitarbeiter-Tenant bekommen. Die üblicherweise sonst von MS Teams für den Beitritt von Gästen in Teams vorgesehene Mechanismen waren daher nicht geeignet, da eine individuelle Einladung jedes einzelnen Teilnehmers durch die Dozierenden nicht praktikabel wäre.

Neben diesen technischen Aspekten sollte beim Einsatz in der Lehre auch Klarheit in der Vielzahl der krisenbedingten Änderungen über den jeweiligen Ort der einzelnen Lehrveranstaltungen bestehen.

Beim parallelen Einsatz von MS Teams neben ILIAS in der Lehre ist dies besonders wichtig. Daher wurde in den Anleitungen grundsätzlich empfohlen, zu jeder Lehrveranstaltung immer als primären Ausgangspunkt den

automatisch aus dem Vorlesungsverzeichnis verlinkten ILIAS-Kurs zu verwenden. Die Beitrittslinks wurden in der Regel in diesen Kursen an die Studierenden weitergeben.



Abbildung 2: Verwaltung eines Teams über connect.studium.kit.edu/teams

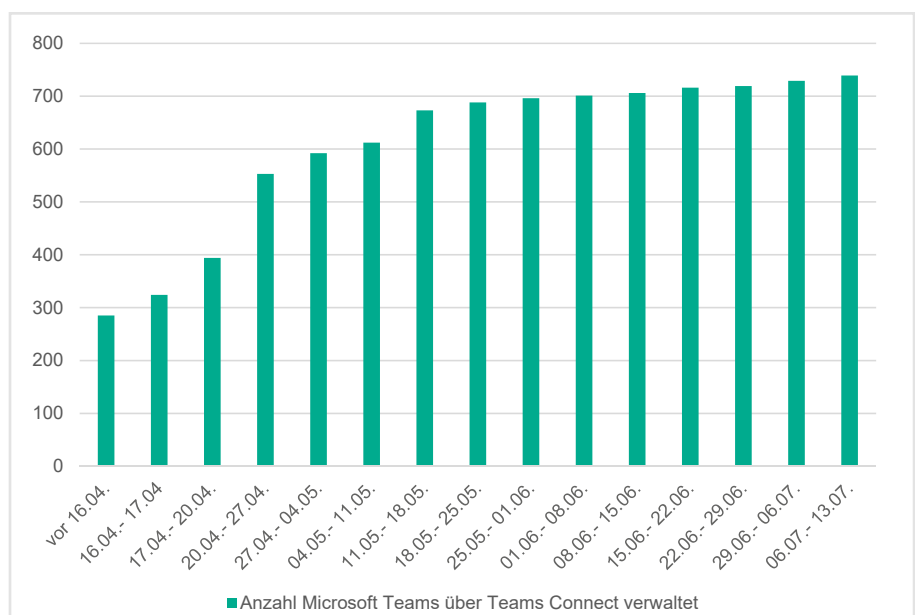


Abbildung 1: Der Einsatz von Microsoft Teams fand schnell Einzug in die Lehre

## Service Support zum Semesterstart

Horst Westergom

Die Service Support-Einheit SCC-Servicedesk (SD) hatte zum Start ins Online-Sommersemester alle Hände voll zu tun. Ohne interne Unterstützung wäre das SD 24\*7 beschäftigt gewesen, doch dank der vielen helfenden Hände aus anderen SCC-Abteilungen konnte der Ansturm von First Level Support-Anfragen gut bewältigt werden.

Für den Support der Konferenzanwendung MS Teams wurde kurzfristig ein abteilungsübergreifendes Team zusammengestellt, das sich aus allen SD-Angehörigen sowie sechs weiteren Personen aus dem SCC zusammensetzte. Dieses Team hat von Mitte April bis Ende Juni 2020 für die Unterstützung bei Fragen rund um MS Teams zusammengearbeitet und konnte in dieser Zeit sehr vielen Anfragenden helfen. Alleine in den ersten Wochen

des Sommersemesters liefen zwischen 1300 und 1500 Anfragen wöchentlich ein.

Ein weiteres Tool, das für die Online-Lehre am KIT eingesetzt wird, ist die Videokonferenzsoftware Zoom. Hier wurde der SD von einem SCC-Kollegen in die grundlegend notwendigen Support-Informationen zu dem Dienst eingewiesen. Zu Zoom gab es zu Beginn des Semesters für den First Level Support relativ wenig Anfragen; im Vergleich zu MS Teams waren es maximal 5%.

Zu dem vom SCC für die Durchführung von Videokonferenzen für mündliche Prüfungen am KIT eingeführten Dienst auf der Basis von Jitsi Meet, sind verglichen dazu fast keine Anfragen eingegangen.



## Vorbereitung und Unterstützung der Lehrenden

Janina Ballach, Andreas Sexauer<sup>1</sup>

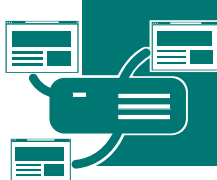
Die schnelle Ausbreitung von Corona in Deutschland hatte auch Einfluss auf die Gestaltung der Lehre im Sommersemester 2020 am KIT. Vorlesungen, Übungen und Tutorien mussten innerhalb kürzester Zeit auf den ausschließlich digitalen Betrieb umgestellt werden. Um allen Lehrenden eine bestmögliche Unterstützung bei der Einarbeitung in die neuen digitalen Lehrszenarien und Tools zu ermöglichen, wurden die zentralen Corona-Info-Seiten unter [www.zml.kit.edu/corona](http://www.zml.kit.edu/corona) geschaffen. Ziel war es, konsistente und verständliche Informationen für alle Lehrenden am KIT zur Verfügung zu stellen – sowohl für Lehrende mit Vorerfahrung als auch für Neulinge in der digitalen Lehre.

Die Info-Materialien entstanden dabei innerhalb kürzester Zeit speziell für diese Zwecke. Durch eine agile Vorgehensweise und frühzeitige Reaktion des Teams sowie eine umfassende Priorisierungsstrategie gelang es, die Materialien sehr schnell zu erstellen und so den Lehrenden frühzeitig Informationen an die Hand zu geben.

So entstanden auf sieben Unterseiten 28 Anleitungen zum Ansehen oder Drucken, davon 13 Anleitungen speziell für die Nutzung von ILIAS, dem Lernmanagementsystem des KIT. Ein Großteil der Anleitungen steht sowohl auf Deutsch als auch für internationale Lehrende auf Englisch zur Verfügung. Neben

Anleitungen finden Lehrende auf den Corona-Info-Seiten auch Hinweise und Tipps zur didaktischen Gestaltung ihrer Online-Lehre. Weiterführende Informationen und hilfreiche Angebote anderer Einrichtungen sind ebenfalls verlinkt. Im Rahmen der Einführung von Microsoft Teams und Zoom am KIT (S. 12 u. 14) veranstaltete das ZML zwei Webinare, deren Aufzeichnungen ebenfalls auf den Seiten zu finden sind. So können sich Lehrende jederzeit selbstständig mit den Programmen vertraut machen. Für Rückfragen und Probleme wurde die E-Mail-Adresse [elarning@studium.kit.edu](mailto:elarning@studium.kit.edu) eingerichtet, über die Lehrende Kontakt zu Experten der digitalen Lehre aufnehmen können.

<sup>1</sup> beide Zentrum für Mediales Lernen (ZML)



## Einführung von Zoom

Als ein wichtiges Szenario für die Online-Lehre wurde „Große Lehrveranstaltungen mit Live-Interaktion (synchron)“ identifiziert. Zu den Anforderungen an eine geeignete Plattform gehörten insbesondere die Skalierbarkeit auch für mehr als 300 Teilnehmer, umfangreiche Moderationsfunktionen und nach Möglichkeit die datensparsame Nutzung durch Studierende ohne eigenes Online-Konto. Diese Anforderungen kann die Plattform Microsoft Teams, die für Lehrveranstaltungen mit kleinen Gruppen genutzt wird, nicht vollständig erfüllen, und auch die Betreiber des bislang am KIT genutzten Videokonferenzdienstes DFNconf signalisierten frühzeitig, dass ein Einsatz für Vorlesungen nicht gewährleistet werden könnte.

*Patrick von der Hagen, Andreas Sexauer<sup>1</sup>*

Vor dem Hintergrund fehlender eigener Betriebserfahrungen und knapper Kapazitäten hat die Expertengruppe für die Online-Lehre frühzeitig empfohlen, nach Möglichkeit kurzfristig Cloud-Dienste zu beschaffen. Unter Abwägung verschiedener Kriterien wurde schließlich die Beschaffung von Zoom beschlossen. Neben einer hohen Skalierbarkeit, die Zoom schon im März weltweit unter Beweis stellen musste, waren die integrierte und einfach zu nutzende Aufzeichnungsmöglichkeit sowie positive Erfahrungsberichte von Lehrenden wichtige Entscheidungsgründe für Zoom.

In kürzester Zeit stimmten sich die Expertengruppe und die Organisationseinheiten Digital Office, Datenschutz und Einkauf mit dem Vizepräsidenten für

Lehre und Akademische Angelegenheiten ab, so dass Angebote eingeholt und Zoom-Lizenzen für die Nutzung in der Lehre beschafft werden konnten.

Durch die Beschränkung auf die Nutzung in der Lehre musste für die Zuteilung der Lizenzen ein pragmatisches, schnelles Vorgehen gefunden werden. Die Lehrenden wurden im Wesentlichen über das Campus-Management-System des KIT identifiziert, es wurde aber auch über die SCC-Gruppenverwaltung ein Weg geschaffen, über den die einzelnen Fakultäten dezentral weitere Lizenzen vergeben können; das ZML übernahm diese Rolle für alle zentralen Einrichtungen.

In enger Absprache mit den Datenschutz- und Informationssicherheitsbeauftragten wurden die Zoom-Voreinstellungen

geprüft, auf für das KIT sinnvolle und datensparsame Vorgaben eingeschränkt sowie organisatorische Regelungen vorgenommen. Deshalb waren am KIT Phänomene wie „Zoom-Bombing“ zu keinem Zeitpunkt ein Problem.

Bei Zoom konnte in einer ersten Einschätzung festgestellt werden, dass eine Anbindung an das IDM und an Shibboleth für SSO gemäß Dokumentation grundsätzlich möglich sein sollte. Nachdem der Zugang zu Zoom bereits am 30. März etabliert wurde, hat das KIT unverzüglich die Einrichtung der nötigen Schnittstellen beantragt, um die Anbindung schnellstmöglich durchzuführen und den Lehrenden frühzeitig zur Semestervorbereitung einen Zugang zu ermöglichen.



Abbildung 1: Zoom kann mit unterschiedlichen Endgeräten genutzt werden. Der Login für Lehrende erfolgt mit dem KIT-Account.



Leider war der Support bei Zoom überlastet und die heiß ersehnten Schnittstellen wurden erst am 14. April um 15:16 Uhr bereitgestellt. Die Anbindung an das SSO (Shibboleth) war kurzfristig möglich, und schon drei Stunden später um 18:24 Uhr konnten schließlich die Dozierenden informiert werden und mit der Nutzung von Zoom beginnen. Zu diesem Zeitpunkt war allerdings nur ein halbautomatischer Import der IDM-Daten über CSV-Dateien möglich, die Automatisierung über eine SCIM-Schnittstelle erfolgte erst nach Semesterbeginn.

Damit blieben den Dozierenden bis Semesterbeginn leider nur drei Arbeitstage, um Erfahrungen mit der neuen Plattform zu sammeln. Allerdings hatte das Zentrum für Mediales Lernen (ZML) die Zeit genutzt, um die Lehrenden durch Materialien, Anleitungen und ein Online-Seminar sowohl inhaltlich als auch didaktisch vorzubereiten und zu unterstützen. Es standen zwei Anleitungen spezifisch zugeschnitten auf den Einsatz in großen Vorlesungen oder in interaktiven Seminaren bereit. Im Online-Seminar wurden die Möglichkeiten mit Zeichenwerkzeugen in der Präsentation, Breakout-Rooms und der gezielten Moderation vertieft.

Lehrende mit einer zugeteilten Lizenz mussten (analog zum Verfahren bei MS Teams) zunächst der Nutzung von Zoom im Self-Service-Portal des SCC zustimmen. Da für Studierende die Nutzung

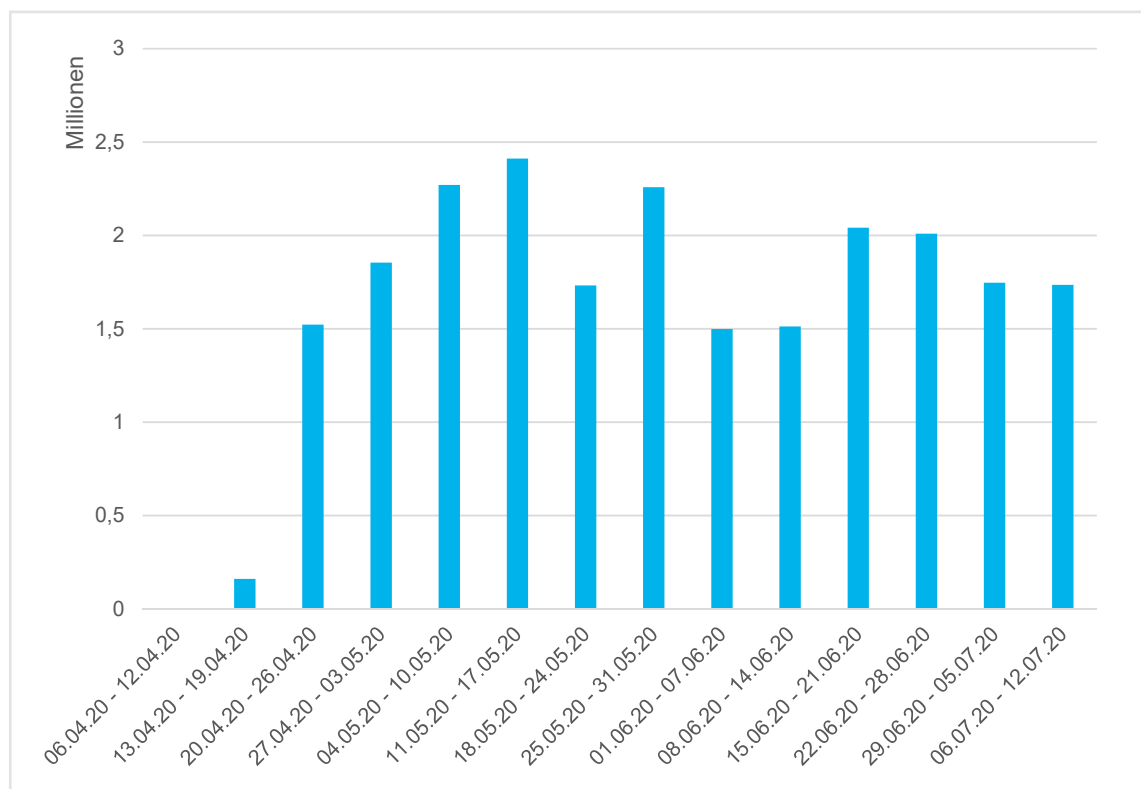


Abbildung 2: Anzahl der Meeting-Minuten über Zoom durch Nutzerinnen und Nutzer des KIT

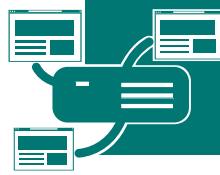
ohne eigenes Konto bei Zoom und mit Pseudonym vereinbart wurde, war für diese eine Zustimmung nicht erforderlich, sie wurden aber über eine eigene Webseite mit Hinweisen, insbesondere zum Datenschutz, informiert.

Durch den seit Beginn der Corona-Pandemie stark angestiegenen Einsatz von Zoom weltweit nahm auch die öffentliche Kritik an Zoom zu, denn es wurden vermehrt Sicherheits- und Datenschutzbedenken gegenüber dem US-amerikanischen Dienstleister geäußert. Diese Kritik wurde offen über die Webseite „Zoom am KIT“ dokumentiert und für das KIT eingeordnet.

Zoom hat im Lauf der Zeit schnell und geeignet auf Kritik und Sicherheitslücken reagiert, so dass z.B. neue Voreinstellungen definiert und den Administratoren des KIT neue Konfigurationsmöglichkeiten eröffnet werden konnten. Diese wurden durchgehend beobachtet und auf die Belange des KIT hin überprüft. Beispielsweise hat das KIT zum ersten möglichen

Zeitpunkt eine Verlagerung der KIT-Zoom Nutzung auf ausschließlich europäische Server veranlasst.

Im Laufe des Semesters konnte eine kontinuierliche Nutzung von Zoom als Werkzeug für Vorlesungen beobachtet werden (vgl. Abbildung 2). Die Rückmeldungen aus der Lehre sind überwiegend positiv. Dennoch bleibt die Kritik an der Cloud-Nutzung. Daher hat sich das SCC frühzeitig mit alternativen Möglichkeiten auseinandergesetzt und beispielsweise eine Online-Prüfungsumgebung mit Jitsi Meet On-Premise etabliert (S. 20). Parallel werden auch weitere On-Premise-Alternativen evaluiert und die nötigen Betriebs Erfahrungen als Grundlage für zukünftige Entscheidungen aufgebaut.



## Wissenwertes zum Opencast-Projekt im Corona-Sommersemester 2020

Mit einem im Oktober 2018 gestarteten Projekt sollte die Hörsaalausstattung am KIT verbessert und ein System zur automatisierten Vorlesungsaufzeichnung aufgebaut werden. Der im Projekt erarbeitete Prototyp bildete glücklicherweise ein solides Fundament für den produktiven Einsatz in der Online-Lehre, so dass daraus bei Ausbruch der Corona-Krise in knapp vier Wochen eine 20 Server umfassende Lösung aufgebaut und betrieben werden konnte.

Uwe Dierolf<sup>1</sup>

### Hintergrund

Im Oktober 2018 wurde am KIT ein Vorlesungsaufzeichnungsprojekt ins Leben gerufen. Es sollte die Hörsaalausstattung verbessert und ein System zur automatisierten Vorlesungsaufzeichnung aufgebaut werden. Die zumeist als Screenshot vorliegenden Videos sollten von einer zentralen Infrastrukturlösung namens Opencast verarbeitet und in die E-Learning-Plattform ILIAS des KIT integriert werden können.

### Vorlesungsvideos im Coronasemester

Die im Verlauf des Projekts gewonnenen Kenntnisse zur Systemarchitektur bildeten ein solides Fundament, so dass bei Ausbruch der Corona-Krise in knapp vier Wochen aus dem entwickelten Opencast-Prototyp eine 20 Server umfassende Lösung vom 5-köpfigen Team der BIB, bestehend aus Uwe Dierolf, Lukasz Koszyk, Tobias Kurze, Andreas Reichert und Udo Willke, in Zusammenarbeit mit dem SCC aufgebaut werden konnte.

Dies wurde seitens der BIB komplett aus dem Homeoffice mit Hilfe der Telko-Lösung Mumble und MS Teams (fürs Screensharing) durchgeführt. Fast jeden Morgen fand ein ausführliches Briefing per Telko statt, in dem die Arbeitspakete verteilt wurden.

### Übersicht über die realisierte Lösung

Durch den Landesdienst bwCloud wurde die Basisinfrastruktur von 20 virtuellen Servern (VM) bereitgestellt. Die einzelnen VMs wurden mit der Automatisierungssoftware Ansible-Playbooks quasi wie von Geisterhand aufgesetzt und der jeweiligen Rolle entsprechend konfiguriert:

- Admin: Zentrale Rolle, die der Steuerung der Opencastkomponenten dient
- Worker: Übernehmen die rechenintensiven Aufgaben wie z.B. Videotranskodierung
- Delivery: Stellen die Videos für das Web bereit

Außerdem eingerichtet wurden Datenbank, Nachrichtenwarteschlangen und Speicherbackends (NFS).

Um besser einschätzen zu können, wie das aufgebaute System skaliert, und um dessen Grenzen zu ermitteln, wurden Lasttests mit teilweise mehr als 10.000 simultanen Zugriffen mit **Loader.io** durchgeführt und damit Worst-Case-Szenarien simuliert.

Aus den daraus gewonnen Erkenntnissen wurde der Beschluss gefasst, die Videos statt mit Full HD- nur in 720p-Auflösung auszuliefern.

Für die Integration in die E-Learning-Plattform ILIAS wurde ein Opencast Plugin genutzt. Das Plugin ist eine Open-Source-Lösung, die von einer Schweizer Firma entwickelt wird. Ein bei intensiver paralleler Nutzung vermuteter Flaschenhals erforderte jedoch die Anpassung des

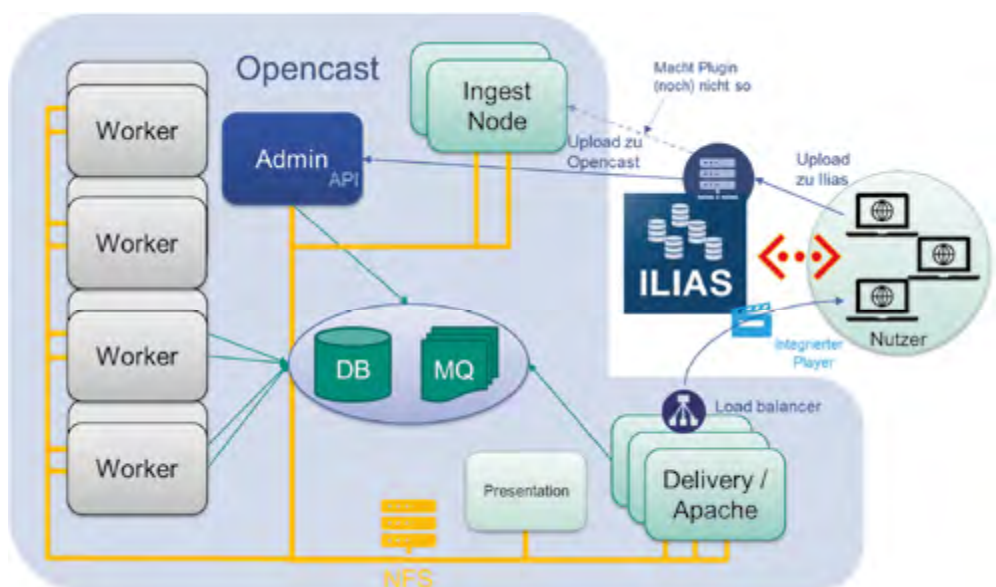


Abbildung 1: Systemarchitektur der Vorlesungsaufzeichnungsplattform

Plugins, welche in Auftrag gegeben und zeitnah fertiggestellt wurde.

Das Plugin erlaubt es, via ILIAS den Opencast-Workflow so anzupassen, dass z.B. Full-HD-Videos nicht zwingend auf 720p heruntergerechnet werden. Das kann bei Kamera-Aufnahmen wichtig sein, wenn die Bilder ein hohes Maß an Detailreichtum aufweisen sollen.

Das transkodierte Videomaterial steht Studierenden in ILIAS durch einen im Plugin eingebauten Player zur Verfügung. Der Player lädt die Videos dabei stückchenweise von den Delivery Nodes nach (progressive download). Beim eigentlichen Download ist Opencast nicht involviert – dieser wird ausschließlich durch Apache-Webserver realisiert. Die Absicherung erfolgt hierbei durch signierte URLs mit begrenzter Gültigkeit.

### Details zu Speicherbackend und Lastverteilung

NFS zur zentralen Datenhaltung ließ Engpässe bei starker Nutzung der Videos erwarten. Geplant war daher, den in der bwCloud selbst gehosteten NFS-Server durch eine Anbindung an die SCC-Fileservices zu ersetzen.

Außerdem sollte zur Lastverteilung von möglicherweise mehr als 10.000 gleichzeitigen Videoabrufen der Lastverteiler F5 des SCC zum Einsatz kommen. Die Integration dieser Komponenten in bwCloud erwies sich aber auf Grund der Komplexität der Netzwerk-Konfiguration als zu aufwändig und zu riskant, da hier Abhängigkeiten von zu vielen Personen und Abteilungen entstanden wären. Daher verzichtete die BIB-IT auf den Einsatz zugunsten simplerer Lösungen – z.B. mittels Apache-Redirects.

### Will it work or does it b(l)end?

Der erste spannende Moment war die Woche vor dem Semesterbeginn. Kann das Gespann bestehend aus ILIAS und Opencast die vielen hundert Video-

Uploads pro Tag störungsfrei überstehen und können die Videos von den Worker-Servern zeitnah verarbeitet werden?

Doch das gerade einmal vier Wochen alte System hielt der Belastung stand. Auch den zweiten Ansturm am Beginn des Semesters haben beide Systeme dank ihrer Skalierbarkeit reibungslos überstanden.

### Kommunikation und Statistiken

Wöchentlich fanden mindestens einmal MIK-Expertengruppen-Meetings statt, bei denen auch das Opencast Team vertreten war. Bei diesen Treffen wurden per Jupyter-Notebook mit Hilfe von Python-Skripten erstellte Auswertungen und aussagekräftigen Statistiken (Gra-

### Briefing der Dozenten und Support

Es wurde schnell erkannt, dass es für ein gutes Funktionieren des Systems sehr wichtig ist, die Dozenten bei der Erstellung der Videos zu unterstützen und zu beraten. Entsprechende Anleitungen wurden in Zusammenarbeit mit dem ZML erstellt. Eine Empfehlung ist, dass Videos eher kurz gehalten und nicht zu hohe Auflösungen gewählt werden sollten. Erfreulicherweise haben die Lehrenden überwiegend kurze Videos (75% aller Videos sind knapp 30 Minuten lang) mit einer Auflösung von 1280x720 hochgeladen. Diese waren meist auch schon vorab sinnvoll komprimiert, was dazu beitrug, die Anzahl fehlgeschlagener Uploads zu

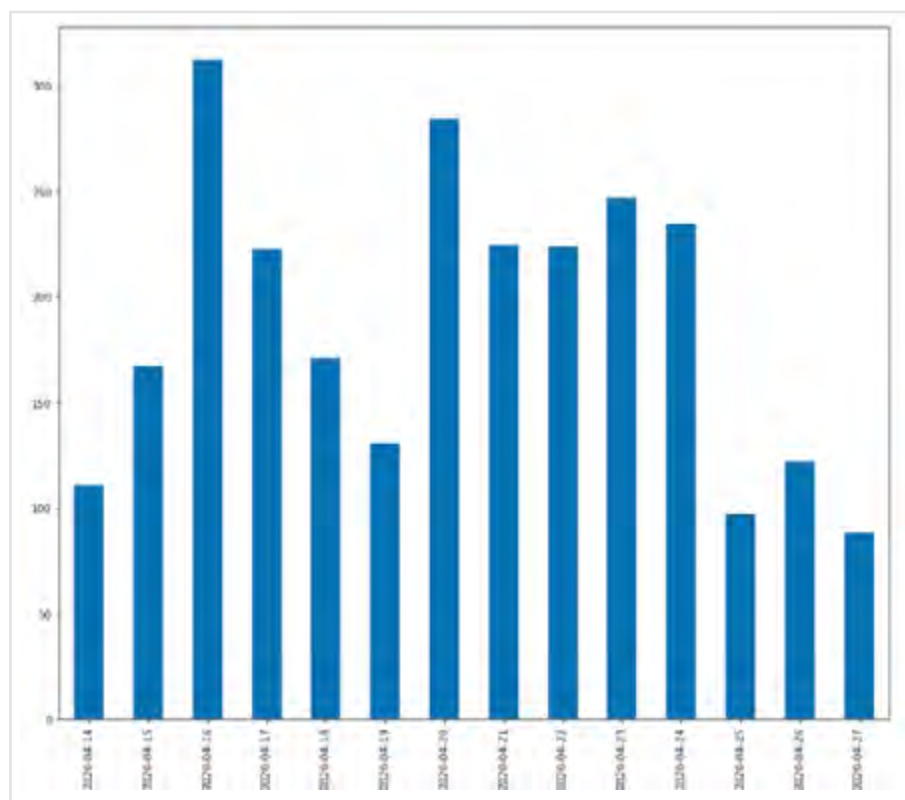


Abbildung 2: Anzahl der täglich hochgeladenen Videos in der Anfangszeit des Semesters

phen und Heatmaps) präsentiert und die Nutzung von Opencast entsprechend dokumentiert.

Mit Stand 1.7.2020 befinden sich ca. 14300 Videos mit einer Gesamtlänge von fast 5000 Stunden in Opencast und sind über ILIAS abrufbar.

reduzieren (Abbildung 3). Die Fehlerquote der Uploads liegt stabil im niedrigen einstelligen Bereich. Das Medienzentrum der BIB konnte den Dozierenden in allen ca. 90 Fällen, bei denen der Upload scheiterte oder das Rendern zu Fehlern führte, umgehend helfen.



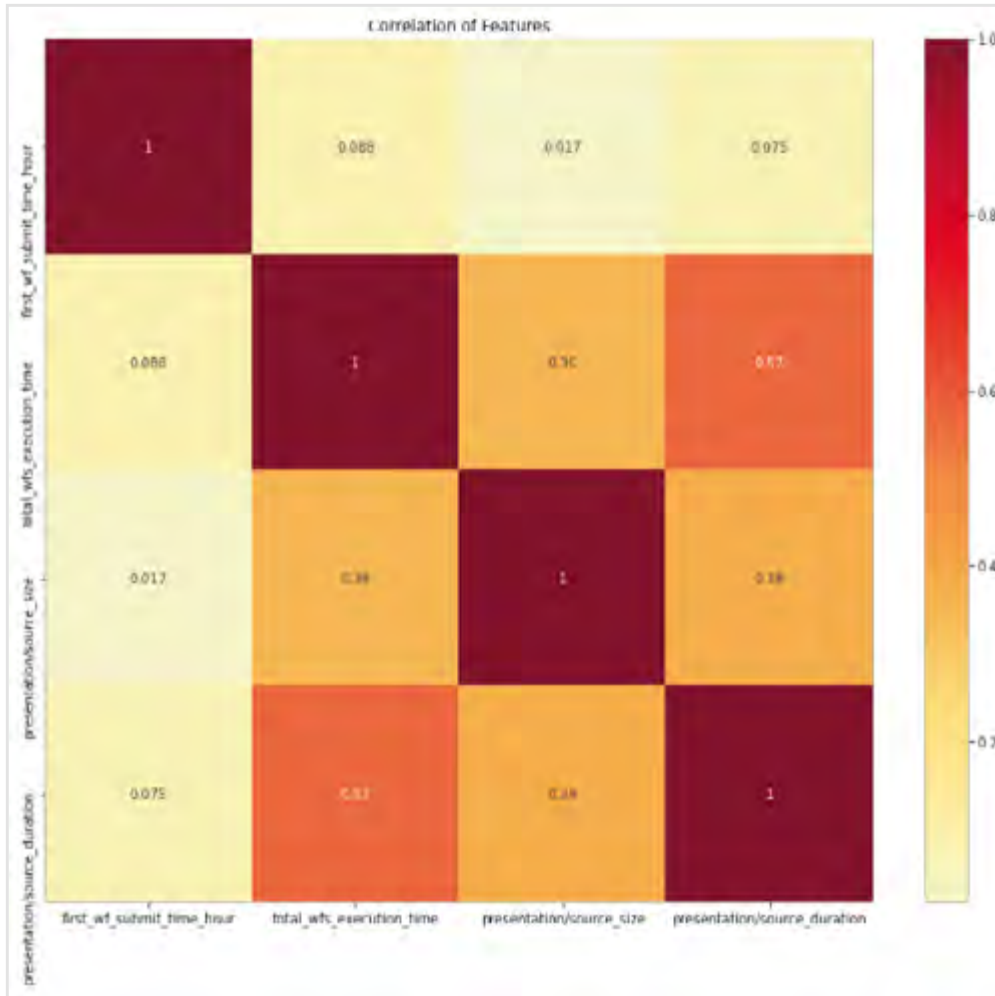


Abbildung 3: Korrelation unterschiedlicher Aspekte. Gut erkennbar ist, dass die Verarbeitungsdauer eines Videos im Wesentlichen von der Länge und nicht von der Kompression bzw. Größe des Ausgangsmaterials abhängt.

### Hilfreiche Tools

Als Monitoring-Tool wird Zabbix eingesetzt, welches Dashboards und Übersichtsscreens bietet. Zusätzlich kann Zabbix als Datenquelle in Grafana zur Darstellung von responsiven Analyse-Web-Charts verwendet werden. Der Zugriff auf die von den Opencast-Entwicklern eingesetzte MariaDB erfolgt via DBeaver. Ein sehr nützliches Tool, das auch ssh-Tunneling erlaubt.

### Ausblick und Fazit

Der Produktivbetrieb der Lösung bringt u.a. die Aspekte Monitoring und Ausbau, Change Management und Know-how-Transfer in den Vordergrund. Das

Opencast-Team beobachtet das System auf kritische Situationen hin, wie z.B. Lastspitzen, Netzausfälle oder Engpässe. Mit Hilfe der o.g. Tools kann schnell analysiert werden, ob nur einzelne Komponenten oder das Gesamtsystem betroffen sind und wie darauf reagiert werden muss.

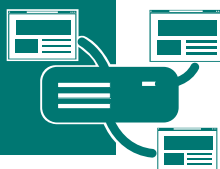
Damit eine große Anzahl an Zugriffen auf die Vorlesungsvideos bewältigt werden kann, stehen acht Delivery-Nodes zur Auslieferung der Videos bereit. Weitere Nodes können sehr einfach hinzukonfiguriert werden. Um jegliche Art von Änderungen sicher testen zu können, wurde ein dem Produktivsystem äquivalentes Testsystem aufgesetzt.

Nicht zuletzt muss das Know-how zum sehr umfangreichen Technology-Stack (von A wie Ansible und Apache über Docker, Elasticsearch, ffmpeg, Grafana, ILIAS, MariaDB, Nginx, Openstack bis Z wie Zabbix) im Betriebsteam verteilt werden, um auch während Abwesenheitszeiten einen stabilen Produktivbetrieb aufrecht zu erhalten.

Als Fazit lässt sich sagen, dass die Zusammenarbeit zwischen SCC, BIB und dem Zentrum für Mediales Lernen (ZML) hervorragend funktioniert hat und sich Homeoffice nie negativ ausgewirkt sondern als beflügelnd erwiesen hat. Auch das Teamwork in der 5-köpfigen BIB-IT war Dank der Tool-Unterstützung, aber vor allem wegen der sehr hohen Motivation und des überragenden Engagements aller Beteiligten sehr konstruktiv. Auch den Familien und Kindern gebührt ein großer Dank, denn Eltern, die im Homeoffice stundenlang am Tag "telkonieren" und sich konzentrieren müssen, sind für alle nicht immer einfach.

### Things worth knowing about the Opencast project in the Corona summer semester 2020

In October 2018, a project was launched to improve the lecture hall equipment at KIT and to set up a system for automated lecture recording. Fortunately, the prototype developed in the project formed a solid basis for productive use in online teaching, so that a solution including 20 servers could be set up and operated in just under four weeks when the corona crisis broke out.



## Mit ILIAS fit für die Online-Lehre

Die Expertengruppe Online-Lehre hat beschlossen, für das Corona-Sommersemester vor allem das Lernmanagementsystem ILIAS zu nutzen, um den Studierenden eine zentrale Informationsplattform zu bieten, wie die Veranstaltungen in dem komplett veränderten Semester ablaufen werden. Für Vorlesungsaufzeichnungen wurde die bereits existierende OpenCast-Infrastruktur der Bibliothek genutzt und in ILIAS über das OpenCast-Plugin integriert (S. 16). Diese Anforderungen verbunden mit dem engen Zeitplan bis zum Start der Vorlesungszeit stellten auch das Betriebsteam der Informationssysteme für Studium und Lehre vor einige Herausforderungen.

Philip Hoyer

Bereits in den Präsenz-Semestern waren die Kapazitäten von ILIAS – insbesondere zum Vorlesungsbeginn – ausgeschöpft, und die Systeme haben unter starker Last operiert, so dass für das Online-Semester entsprechend gehandelt werden musste.

Die Abteilung ISL hat daher die Ressourcen und Kapazitäten der Lernplattform

ILIAS am KIT etwa um den Faktor drei erweitert und ist dabei bis an die Grenzen der SCC-Virtualisierungsumgebung (VMware ESXi) gegangen. Die ILIAS-Plattform ist im Kern eine klassische Web-PHP-Anwendung mit einem entsprechenden Schichtenaufbau. Die am SCC eingesetzten Load Balancer F5 BIG-IP verteilen die Anfragen der Nutzer an die Webserver,

welche wiederum den Datenbankserver und den Fileserver aufrufen (Abbildung 1).

Bis zum Wintersemester 2019/20 waren für ILIAS vier Webserver im Einsatz, die mit dem Webserver Apache 2 und integriertem PHP 7 (mod\_php) liefen. Diese wurden außer Betrieb genommen und durch 12 vollständig neue virtuelle Webserver mit der aktuellen Version des Linux-Betriebssystems Debian ersetzt. Außerdem wurden die neuen

Apache-Webserver mit den Softwaremodulen mpm\_event, PHP-FPM und HTTP/2 konfiguriert, um die Auslieferung der ILIAS-Webseiten zu beschleunigen. Die Funktionsfähigkeit und Auslastung der Systeme werden mit der Monitoring-Software Telegraf überwacht, die Telemetriedaten können über Grafana visualisiert werden.

Die ILIAS-Datenbank läuft auf einem MariaDB-Server, dessen Ressourcen ebenfalls verdreifacht wurden. Der Datenbankserver operiert nun mit 24 CPU-Cores und 64 GB RAM. Trotz Virtualisierung belegt der Datenbankserver einen kompletten ESXi-Host und kann ohne intelligente Softwareanpassungen nur durch noch leistungsfähigere Hardware erweitert werden (vertikale Skalierung oder „scale up“). Für das Wintersemester 2020/21 ist statt dessen der Aufbau einer Master/Slave-Replikation mit vorgeschaltetem Proxy (MariaDB MaxScale) anvisiert, um insbesondere die lesenden Anfragen an die Datenbank, die den überwiegenden Teil der Anfragen ausmachen, auf mehrere Datenbankserver zu verteilen (horizontale Skalierung oder „scale out“). Damit kann die Last auf mehrere Server verteilt werden.

Für den Abruf der Vorlesungsvideos wurde in ILIAS das OpenCast-Plugin konfiguriert und an die Systeme der KIT-Bibliothek angebunden. Damit können Dozenten Aufzeichnungen komfortabel über ILIAS hochladen und den Studierenden zur Verfügung stellen. Ebenso können Studierende die Aufzeichnungen über ILIAS anschauen, der Videostream läuft dabei aber

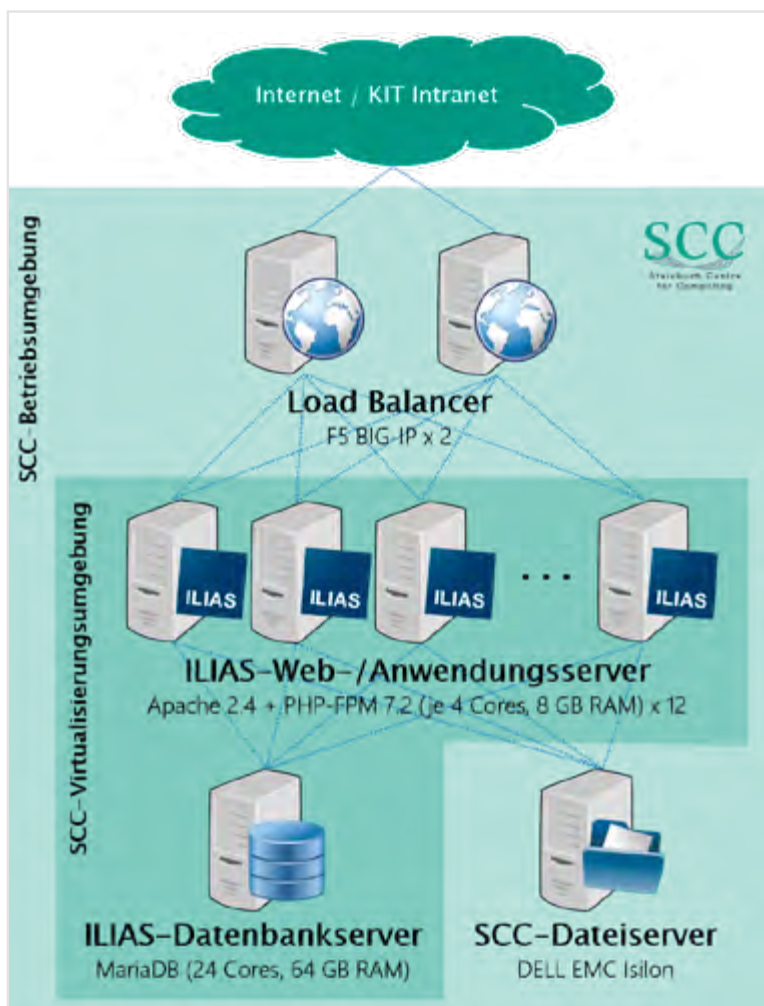


Abbildung 1: Systemarchitektur der ILIAS-Lernplattform

über die OpenCast-Server. Einen potentiellen Performance-Engpass im OpenCast-Plugin konnte die Firma, die die Software entwickelt hat, rechtzeitig überarbeiten und für den Betrieb ausliefern.

Rückblickend waren die vorgenommenen Erweiterungen für das Sommersemester 2020 keinesfalls überdimensioniert, aber auch nicht stark unterdimensioniert. Insbesondere der Beginn der Vorlesungszeit,

der üblicherweise mit sehr hoher Nutzung für ILIAS verbunden ist, hat nach Rückmeldungen der Dozierenden und Studierenden und auch anhand der Telemetriedaten sehr gut funktioniert.



**Mündliche Online-Prüfungen brauchen einen besonders geschützten Raum, dem alle Beteiligten vertrauen können.**

## Online-Prüfungen in Zeiten der Pandemie

Während für Vorlesungen, kleinere Arbeitsgruppen und Seminare die Konferenzsoftware Zoom und Microsoft Teams zur Verfügung standen, musste für mündliche Prüfungen per Videokonferenz eine sichere und höheren Datenschutzerfordernungen genügende Lösung gefunden werden. Das SCC hat daher eine auf der Webanwendung Jitsi Meet basierende Videokonferenzplattform entwickelt. Über ein leicht zu bedienendes Online-Portal können Prüferinnen und Prüfer eigene, am KIT gehostete, Videokonferenzräume buchen.

*Samuel Braun, Peter Krauß, Ulrich Weiß, Alvar Wenzel*

Der Vizepräsident für Lehre und akademische Angelegenheiten, Alexander Wanner, hat Ende März den Wunsch an das SCC getragen, eine Videokonferenzplattform bereitzustellen, die den für Prüfungen bestehenden hohen Datenschutzerfordernungen genügt. Diese sollte deshalb "On-Premises" betrieben werden, das heißt die Software läuft auf Servern des SCC. Die Abteilungsleiter des SCC stellten ein abteilungsübergreifendes

des Team zusammen, um unterschiedliche Expertisen zu bündeln. Ziel war es, möglichst schnell zu agieren und eine datenschutzkonforme und sichere, gleichzeitig aber einfach zu bedienende und betriebssystemunabhängige Videokonferenzlösung zu etablieren. Zeitnah folgten erste Projekt- bzw. Koordinierungstreffen, an denen mit Vertretern aus den Abteilungen AMA, NET, SCS, SDM und SYS ein breitgefächertes Spektrum an Experten-

wissen zusammen kam. Zusätzlich konnte ein HiWi sein in einem Seminar erarbeitetes Wissen über das Cloud-Betriebssystem OpenStack einbringen.

### Bedarfs- und Ressourcenanalyse

Während der ersten virtuellen Zusammenkünfte wurden unterschiedliche Videokonferenz-Softwarelösungen evaluiert. Die Open-Source-Programmpakete BigBlueButton und Jitsi kamen in die engere Auswahl. Da die Anzahl an gleichzeitig benötigten Konferenzräumen unbekannt war, musste auf eine gute Skalierbarkeit der avisierten Lösung geachtet werden. Die Art und Weise, wie die Terminplanung und Konferenzraum-Zuteilung zu erfolgen hatte, stand ebenso wie die Wahl der zu verwendenden Compute-Ressourcen zur Diskussion. Ein wichtiges Augenmerk wurde auf Sicherheitsaspekte gelegt.

Das SCC-Team konnte in verschiedenen Gruppen agil voranschreiten und

Abbildung 1 zeigt das Frontend zur Buchung einer Prüfung.

Entwicklungen parallel vorantreiben. Alle waren zu diesem Zeitpunkt selbst im Homeoffice und zudem teilweise in anderen, ähnlich priorisierten Projekten eingebunden.

### Technische Umsetzung und Bedienkonzept

Das nach den Treffen erarbeitete Konzept wurde für gut befunden und die Videokonferenzsoftware Jitsi Meet ausgewählt. Sie verfügt über einen ausreichenden Funktionsumfang, welcher auch in Stresssituationen nicht zu Überforderung führt. Weiterhin werden alle modernen Web-Browser unterstützt, wobei allerdings ein Chrome/Chromium Browser oder die Smartphone-App favorisiert wird.

Als Compute-Ressource für Backend und Middleware wurde die Karlsruher OpenStack Instanz der „bwCloud“ gewählt. Frontend und MySQL-Datenbank laufen ebenfalls auf vom KIT betriebener Hardware. Aus Skalierungs- und Sicherheitsgründen wurden die Ressourcen „On-Demand“ bereitgestellt. Die Granularität wurde auf eine Jitsi-Instanz pro Tag und Prüfer bzw. Prüferin festgelegt. Pro Prüfung wird dabei ein neuer Konferenzraum erstellt, der lediglich den Prüfungsbeteiligten bekannt ist. Zusätzlich wird zum Betreten ein KIT-Konto vorausgesetzt.

### Von der Buchung bis zur Prüfung

Die Planung der Prüfung sollte bereits einige Tage vor der Prüfung erfolgen. Dies ermöglicht einen Testlauf, und eventuell auftretende technische Probleme können vorab gelöst werden. Weiterhin ist es vorteilhaft, allen Prüfungsteilnehmenden die Zugangsdaten vorab zukommen zu lassen. Es wäre auch möglich, noch am Tag der Buchung eine Prüfung abzuhalten, da die benötigte Vorlaufzeit lediglich wenige Minuten beträgt, bis eine neue Instanz mit einem Prüfungsraum eingerichtet und verfügbar ist. Die Buchung erfolgt komfortabel über ein Web-Portal ([vc-exam.scc.kit.edu](http://vc-exam.scc.kit.edu)). Der Zugang zum

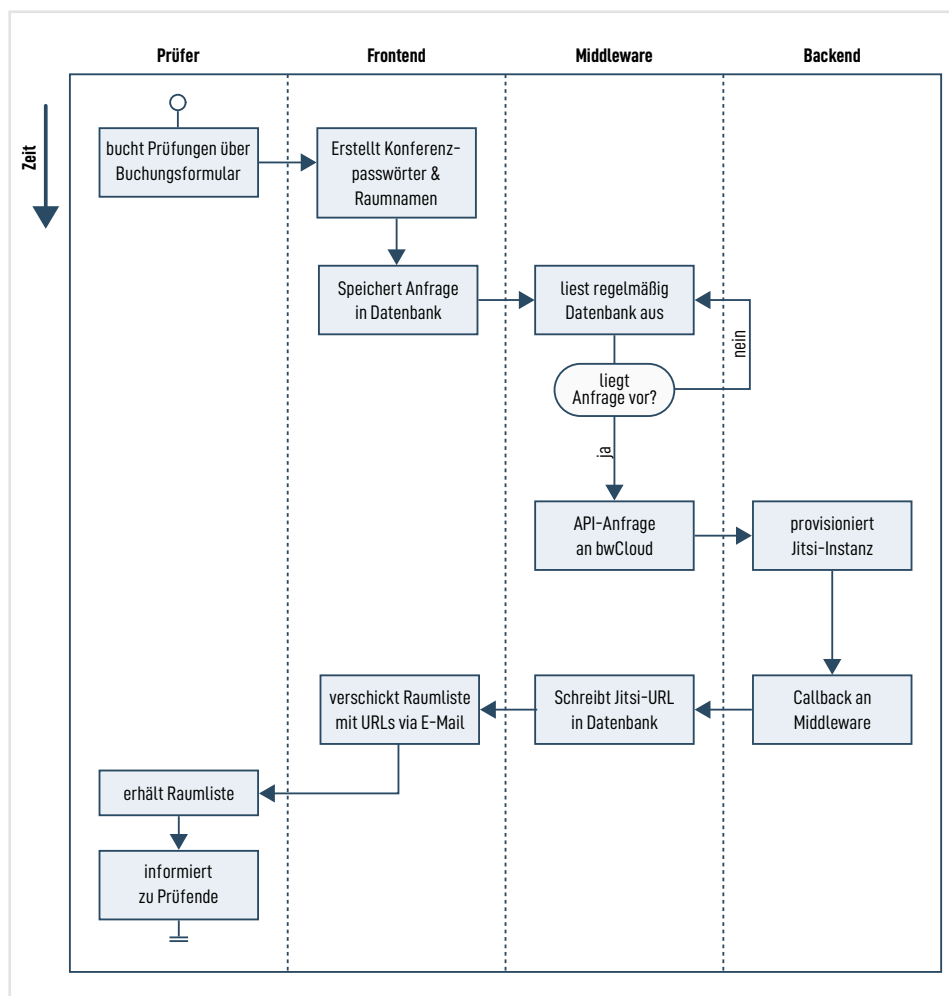


Abbildung 2: Das Schaubild zeigt den vereinfachten Ablauf zum Buchen einer Prüfung mit dem KIT VC-Exam System. Der Prüfer füllt ein Formular aus, worauf hin ein Datenbankeintrag erstellt wird. Auf Basis der Anfrage wird dann eine neue Jitsi-Instanz angelegt und die URL via E-Mail zurückgeschickt.

Buchungssystem ist auf die im Campus-Management-System als Prüfungsberechtigte eingetragenen Personen beschränkt. Die Verwaltung dieser Rolle erfolgt in den Fakultäten.

Sobald die Buchung erfolgt ist, erhält der oder die Prüfende eine Bestätigungs-Mail mit den wesentlichen Informationen, unter anderem den Link zum virtuellen Videokonferenzraum, der an den Prüfling weitergegeben werden kann. Sollte der Link außerhalb des Prüfungszeitraums aufgerufen werden, wird an eine Test-Instanz weitergeleitet. Auf diese Art können sich alle Beteiligten vorab mit dem System vertraut machen und eventuelle Probleme mit Audio- und Videogeräten beheben.

### Technische Umsetzung

Der Buchungsworkflow ist in zwei Teilen realisiert: Dem PHP-basierenden Frontend zur Buchung und Durchführung der Prüfungen sowie einer Python-Middleware zum Management der Jitsi-Instanzen. Diese beiden Teile sind über eine MySQL-Datenbank als Basis verbunden.

### Buchungs-Portal als Frontend

Das Frontend ermöglicht einerseits das Erfassen neuer Prüfungen, andererseits können die bereits gebuchten Einträge bearbeitet werden. Die Buchung von Prüfungen erfolgt dabei tagesweise unter Angabe der Anzahl jeweils benötigter Prüfungsräume und der dabei auf Prüferseite beteiligten Personen. Nach Buchung



einer neuen Prüfung werden die wesentlichen Informationen zusätzlich per E-Mail an die eingetragenen Personen versandt, darunter die URLs der später generierten Jitsi-Instanzen sowie die Zugangsdaten zur Öffnung der Prüfungsräume. Diese Informationen finden sich auch in der Liste der bereits gebuchten Prüfungstage im Portal. Bei Bedarf können weitere Prüfungsräume gebucht werden.

### Middleware

Die Middleware steuert die Provisionierung der Hardware-Ressourcen für die vorhandenen Buchungen. Jitsi-Server werden nur für die Prüfungen erstellt, die in den nächsten 48 Stunden stattfinden sollen. Jeder Prüfer erhält so einen „eigenen“ Server mit mehreren, passwortgeschützten Prüfungsräumen. Die Middleware nutzt hierfür die OpenStack-API und ein Ubuntu-Basistemplate. Zunächst wird ein passender Hardware-Server der bw-Cloud-Infrastruktur ausgewählt, und zwar so, dass auf der verwendeten Hardware möglichst wenige Jitsi-Server gleichzeitig genutzt werden. Dies dient einer bestmöglichen Lastverteilung auf den in der bwCloud reservierten Ressourcen. Danach wird eine neue virtuelle Maschine erstellt

und das Basistemplate bezüglich der Umgebung angepasst. Nötige Einstellungen sind hier beispielsweise Hostname, Netzwerkadresse oder Firewall-Einstellungen. Dann erfolgt die automatische Installation der Jitsi-Software. Der gesamte Provisionierungsvorgang wird von der Middleware überwacht. Bei eventuell auftretenden Problemen wird unterschiedlich reagiert. Bei kleineren Problemen, beispielsweise bei fehlerhaften Checksummen heruntergeladener Software-Pakete oder temporärer Nichtverfügbarkeit von Netzwerkressourcen, wird der Vorgang wiederholt. Bei schwerwiegenden Fehlern wird das Administratorenteam per Mail informiert. Nachdem alle Prüfungen erfolgt sind, steuert die Middleware zeitversetzt die Deprovisionierung des Jitsi-Servers.

### Inbetriebnahme

Der erste Prototyp der Plattform war etwa eine Woche nach Projektstart funktionsfähig. In der zweiten und dritten Woche wurde insbesondere an Layout, Bedienkomfort und (Daten-) Sicherheit gearbeitet. So wurde unter anderem ein eigener STUN-Server eingerichtet, um einen vollständigen „On-Premises“-Betrieb sicher zu stellen. Parallel zur technischen

Entwicklungsarbeit wurde eine verständliche Dokumentation für Prüfungspersonal und zu Prüfende erstellt. Zu guter Letzt musste ein Administrationsteam sowie der technische Support über das SCC-Servicedesk organisiert werden. Das Zentrum für Mediales Lernen (ZML) des KIT unterstützt den Dienst methodisch und didaktisch.

### Fazit

Das Projekt „Onlineprüfungen On-Premises“ war für alle Beteiligten arbeitsintensiv. Die abteilungsübergreifende Herangehensweise bot jedem Einzelnen die Gelegenheit, bei einer spannenden Tätigkeit den gesamten Querschnitt des SCC kennen zu lernen. Sehr gut funktioniert hat die Mitwirkung der wissenschaftlichen Hilfskraft an zentraler Stelle des Projekts. Sie konnte ihre Expertise gestaltend einbringen und bekam so Einblicke in ein sehr praxisnahes Projekt.

Seit Inbetriebnahme wurden über 300 Jitsi-Instanzen bereitgestellt, auf denen über 1.600 Prüfungsräume eingerichtet worden sind. Mit Hilfe dieser Plattform konnte der Prüfungsbetrieb des KIT auch in Zeiten der Krise aufrechterhalten werden.

### Online exams in times of pandemic

While the conference software Zoom and Microsoft Teams were available for lectures, smaller working groups and seminars, a secure solution meeting higher data protection requirements had to be found for oral examinations via video conferencing. The SCC therefore developed a video conferencing platform based on the web application Jitsi Meet. Via an easy-to-use online portal, examiners can book their own video conference rooms hosted at KIT.

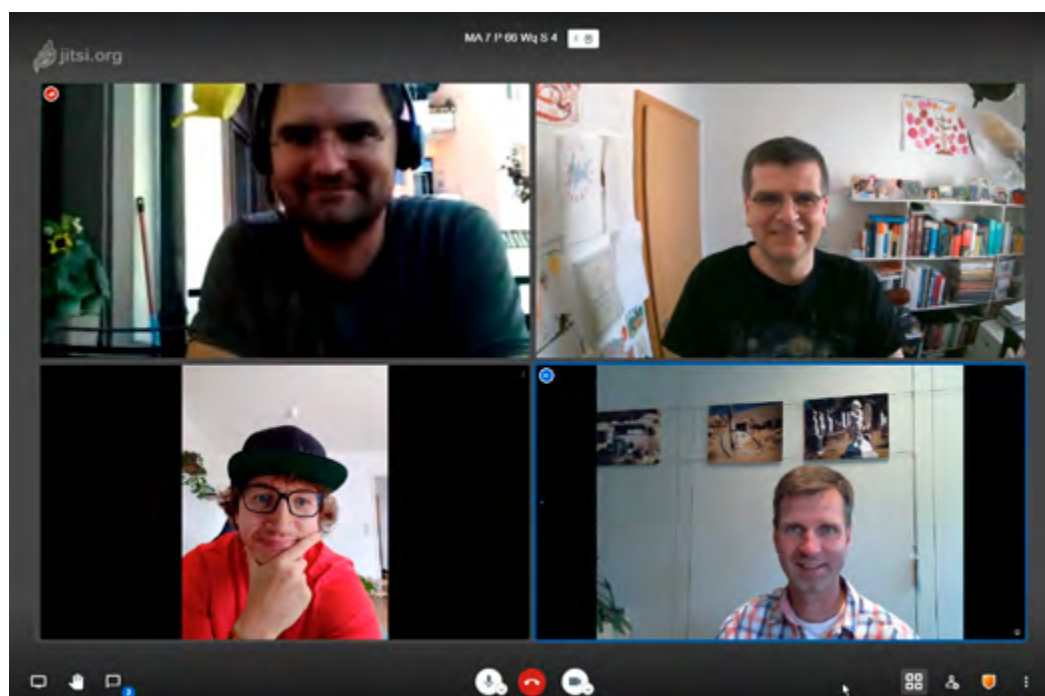


Abbildung 3: Das „VC-Exam“-Team Samuel Braun, Alvar Wenzel, Peter Krauß und Ulrich Weiß (v. links oben n. rechts unten)

# CAS Campus – zentrales Campus-Management-System am KIT

Das Softwareprodukt CAS Campus der CAS Software AG wurde am KIT bereits 2011 als zentrale Campus-Management-Lösung eingeführt. Das System wurde in den Folgejahren kontinuierlich an die Prozesse in Studium und Lehre am KIT angepasst und erweitert. Immer mehr Verfahren sind nun direkt im System abgebildet, so dass aufwändige, zum Teil papierbasierte Prozesse wegfallen und Arbeitsabläufe rein digital quer über die unterschiedlichen Stellen sicher und effizient abgewickelt werden können. In diesem Artikel werden einige Features vorgestellt, die das SCC in den letzten Jahren zur Umsetzung gebracht hat.

Philip Hoyer

Bereits seit 2011 ist CAS Campus das zentrale Campus-Management-System am KIT (campus.kit.edu). War CAS Campus zunächst nur für das Veranstaltungsmanagement im Einsatz, wurden seit 2013 auch erste Studiengänge vom bisherigen Hochschulinformationssystem HIS-POS nach CAS Campus übertragen. Die Studiengänge wurden dabei nicht eins-zu-eins übertragen, sondern vollständig neu abgebildet. Dies geschah einerseits durch die Vorgaben der Rahmenprüfungsordnung, die sich das KIT im Rahmen der Systemakkreditierung für die externe Qualitätssicherung gegeben

hatte, zum anderen auch aus technischen Gründen und wegen der Möglichkeiten der Software. Hierbei zeigten sich schnell Defizite in CAS Campus, die aber das KIT durch entsprechende Weiterentwicklungen zusammen mit der CAS Software AG ausräumen konnte. Im Ergebnis konnte Anfang letzten Jahres der letzte Studiengang nach CAS Campus überführt werden. Somit werden alle, die sich neu am KIT einschreiben, in CAS Campus geführt. Im Folgenden sollen daher einige Features vorgestellt werden, die das SCC in den letzten Jahren zur Umsetzung in CAS Campus beauftragt hat.

## Flexible Auswahlmöglichkeiten für Studierende

Als erster Meilenstein wurde 2017 CAS Campus um flexiblere Wahlmöglichkeiten in den Studiengängen erweitert. Studierende können damit im Rahmen der vorgesehenen Möglichkeiten selbstständig gewünschte Studiengangbestandteile wählen. Davor war die Wahl von Teilleistungen auf Modulebene nicht möglich. Für einige Studiengänge, z.B. Maschinenbau, war dies für eine vollständige Abbildung in CAS Campus jedoch eine notwendige Voraussetzung. Innerhalb

**Wahl (Modulebene)**

✔ Gewählte Leistungspunkte: 12,00 von min. 12,00 (Wahl bei min. LP abgeschlossen)

**Pflichtbestandteile**

Ä	G	Z	Titel	LP	Gewicht	Terminierung (Semester)
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	T-MACII-110035 – Kontinuumsmechanik der Festkörper und Fluide	4,0	1	5
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	T-MACH-110836 – Mathematische Methoden der Kontinuumsmechanik	4,0	1	5

**Kontinuumsmechanik (E)**

Ä	G	Z	Titel	LP	Gewicht	Terminierung (Semester)
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	T-MACH-110362 – Einführung in die Numerische Strömungsmechanik	4,0	1	5
<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	T-MACH-110837 – Einführung in die Finite-Elemente-Methode	4,0	1	5

Geänderte Wahl (Spalte Ä) genehmigen

Ursprüngliche Wahl (Spalte G) genehmigen

Wahl ablehnen und Ursprungszustand (Spalte Z) herstellen

Abbrechen

Abbildung 1: Die Wahl des Studierenden kann durch Verantwortliche abgelehnt bzw. mit oder ohne Änderungen genehmigt werden.

**Anzeigeoptionen**

Umfang:  Alle Daten  
 Nur Veränderungen

Relevanz:  Keine Einschränkungen  
 Nur abbildungsrelevante Daten

	(Version 3)	(Version 4)
▶	<b>Stammdaten</b>	
▶	<b>Wahlkriterien</b>	
▶	<b>Modulversionsaufbau</b>	
▲	<b>Arbeitsaufwand</b>	
☀	Standardwert: siehe einzelne Teilleistungen	Der Arbeitsaufwand im Bachelor of Science beträgt ca. 360 Zeitstunden, wovon etwa 125 Stunden Präsenzzeit darstellen.
▲	<b>Arbeitsaufwand (EN)</b>	
☀	Standardwert: see different bricks	The work load is about 360 hours in the Bachelor of Science program, whereof the presence time is 125 h.
▲	<b>Empfehlungen</b>	
☀	Standardwert: Studierenden der Fachrichtung Maschinenbau, die den Schwerpunkt 13 gewählt haben, empfehlen wir für das Wahlpflichtmodul folgende Lehrveranstaltung: "Wissenschaftliches Programmieren für Ingenieure" (Prof. Gumbusch, Dr.Weygand, Wintersemester) oder "Einführung in Python" (Dr. Weiß, Sommersemester)	Studierenden der Fachrichtung Maschinenbau, die den Schwerpunkt 13 gewählt haben, empfehlen wir für das Wahlpflichtmodul folgende Lehrveranstaltung: "Wissenschaftliches Programmieren für Ingenieure" (Prof. Gumbusch, Dr.Weygand, Wintersemester)

Abbildung 2: Vergleich zwischen zwei Modulversionen.

dieser Softwareerweiterung wurde auch ein digitaler Workflow zur Genehmigung der Auswahl von Studiengangbestandteilen durch Studierende implementiert. In einigen Studiengängen müssen die von Studierenden getroffenen Wahlentscheidungen zunächst von den Fakultäten genehmigt werden, bevor diese belegt werden können. Dies erfolgte bislang durch papierbasierte Prozesse. Mit dem neuen Genehmigungsprozess können Studierende nun selbstständig einen digitalen Wahlantrag für Module oder Teilleistungen einreichen. Der Antrag wird von Wahlgenehmigenden an den Fakultäten geprüft und kann bestätigt, abgelehnt oder geändert werden (siehe Abbildung 1). Änderungen an bereits genehmigten Wahlentscheidungen sind möglich, selbst wenn Teile der Wahl bereits belegt wurden.

### Mehrfachstudiengänge

Ein wichtiger Meilenstein wurde im Jahr 2018 mit der Implementierung der

Mehrfachstudiengänge in CAS Campus erreicht. Am KIT sind Mehrfachstudiengänge insbesondere bei Lehramtsstudiengängen und den neuen „Bachelor/Master of Education“-Studiengängen von Bedeutung. Eine wichtige Anforderung war es, speziell hierfür entsprechende Zeugnisse mit einer Gesamtnote über alle Teilstudiengänge ausweisen zu können. Spezifikation und Beauftragung der dafür benötigten Features wurden in Kooperation mit der Ruhr-Universität Bochum (RUB) durchgeführt, um die dafür notwendigen Aufwände auf mehrere Schultern zu verteilen.

### Verbesserte Workflows und User-Experience

Außerdem wurden an unterschiedlichen Stellen Optimierungen der Software durchgeführt, welche die Nutzerinnen und Nutzer bei verschiedenen Arbeiten mit dem System besser unterstützen. Beispielsweise können Studiengangbestandteile nun dezentral durch berech-

tigte Personen gelöscht werden, wenn diese nicht mehr benötigt werden oder irrtümlich angelegt wurden, und müssen nicht mehr durch administrative Eingriffe von zentraler Seite entfernt werden. Ebenso wurde die Funktion „Modulversionen vergleichen“ optimiert, die es erlaubt, die Änderungen zwischen zwei Versionen eines Moduls komfortabel gegenüberzustellen (siehe Abbildung 2).

### Studienabläufe individuell planen

2019 konnte die individuelle Anpassung von Studienablaufplänen umgesetzt werden. Durch diese Erweiterung können einerseits bei einzelnen Studierenden im vorgesehenen Studienablaufplan konkrete Module

oder Teilleistungen, die eigentlich nicht im vorher definierten Ablauf vorgesehen sind, hinzugefügt werden, und andererseits Pflichtmodule oder Pflichtteilleistungen deaktiviert werden, damit diese nicht absolviert und bestanden werden müssen (siehe Abbildung 3). Auf diese Weise kann das Studium für einzelne Studierende individueller gestaltet und Sonderfälle können direkt im System berücksichtigt werden, die ansonsten über umständliche, papierbasierte Ausnahmeregelungen bearbeitet werden müssten.

### Weitere Features geplant

Noch für das Jahr 2020 sind zwei wesentliche neue Features geplant. Zum einen soll das „Verschieben“ von Modulen und Teilleistungen in einem Studienablaufplan erheblich vereinfacht werden, so dass statt des momentan minutenlangen Umbuchungsprozesses durch die Fakultäten bzw. den Studierendenservice ein einfaches Verschieben mit zwei bis drei Mausklicks in wenigen Sekunden möglich

wird. Möglicherweise kann dieses Feature dann auch den Studierenden zur Verfügung gestellt werden, um Leistungen selbst umzubuchen. Zum anderen soll die sogenannte 4,0-Bescheinigung durch Prüfende digital ausgestellt werden können. Durch eine 4,0-Bescheinigung bestätigt der Prüfende eine Prüfung als bestanden, obwohl die Note noch nicht feststeht. Sofern es sich dabei um die letzte Prüfung in einem Studium handelt, kann die Titel-

führungsbescheinigung den Studierenden direkt ausgestellt werden, und diese sind somit berechtigt, den entsprechenden akademischen Grad zu führen.

#### Ausblick: Digitale Prozesse steigern die Effizienz

Viele papierbasierte Prozesse im Umfeld der Lehre und der Verwaltung von Studierenden konnten durch die beschriebenen

Erweiterungen der Software entfallen. Bislang zentral wahrgenommene Aufgaben, die oftmals zu Verzögerungen in den Prozessen geführt haben, wurden mit den entsprechenden Softwareanpassungen nun an dezentrale Positionen in den KIT-Fakultäten verlagert. Die Folge ist eine schnellere Bearbeitung von Anfragen bei gleichzeitig spürbaren Arbeitsentlastungen und geringerem Kommunikations-Overhead.

**Wahl (Modulebene)**

**✓ Gewählte Leistungspunkte: 12,00 von min. 12,00 (Wahl bei min. LP abgeschlossen)**

**Pflichtbestandteile**

Titel	LP	Gewicht	Terminierung (Semester)
<input type="checkbox"/> T-MACH-110377 – Kontinuumsmechanik der Festkörper und Fluide	3,0	1	5
<input checked="" type="checkbox"/> T-MACH-110836 – Mathematische Methoden der Kontinuumsmechanik	4,0	1	5

**Individuelle Pflichtbestandteile**

Element hinzufügen...

Titel	LP	Gewicht	Terminierung (Semester)
<input checked="" type="checkbox"/> T-INFO-104771 – Visual Computing	3,0	1	5

**Kontinuumsmechanik (E)**

Titel	LP	Gewicht	Terminierung (Semester)
<input checked="" type="checkbox"/> T-MACH-110362 – Einführung in die Numerische Strömungsmechanik	3,0	1	5
<input type="checkbox"/> T-MACH-105320 – Einführung in die Finite-Elemente-Methode	3,0	1	5

**Kontinuumsmechanik (Ü)**

Titel	LP	Gewicht	Terminierung (Semester)
<input checked="" type="checkbox"/> T-MACH-110330 – Übungen zu Einführung in die Finite-Elemente-Methode	1,0	1	5
<input checked="" type="checkbox"/> T-MACH-111033 – Übungen zu Einführung in die Numerische Strömungsmechanik	1,0	1	5
<input type="checkbox"/> T-MACH-110333 – Übungen zu Kontinuumsmechanik der Festkörper und Fluide	1,0	1	5

Abbildung 3: In der Wahlmaske können Pflichtelemente abgewählt und beliebige individuelle Pflichtelemente hinzugefügt werden.

#### CAS Campus – Central Campus Management System at KIT

The software product CAS Campus of CAS Software AG was already introduced at KIT in 2011 as a central campus management solution. In the following years, the system was increasingly adapted and expanded to better match the processes in studies and teaching at KIT. More and more procedures are now mapped directly in the system, so that complex, partly paper-based processes are no longer necessary and workflows can be handled digitally across the different departments in a secure and efficient way. This article presents some features that the SCC has implemented in the last years.

Das SCC wird weiterhin an der Optimierung der Funktionalitäten von CAS Campus arbeiten, sowohl durch Beauftragen der CAS Software AG als auch durch entsprechende Eigenentwicklungen. Ziel war und ist es, die Prozesse in Studium und Lehre weiter zu digitalisieren und alle an diesen Prozessen beteiligten Personen zu entlasten, damit diese sich auf ihre eigentlichen Aufgaben konzentrieren können. Im Rahmen der IV-Governance erfasst das SCC hierzu mit dem Arbeitskreis Informationsversorgung Studium & Lehre (AK IV SuL) die Anforderungen der KIT-Fakultäten, der Studierenden und der zentralen Einrichtungen in einem digitalen Kanban-Board über die Software JIRA. In regelmäßigen Intervallen werden die so gesammelten Anforderungen mit dem Arbeitskreis hinsichtlich ihrer Priorität betrachtet und gemeinsam zur Umsetzung eingeplant.



## Akademische COMSOL Multiphysics Campus-Lizenz für das KIT

COMSOL Multiphysics ist eine Applikation zur Modellierung und Berechnung von Multiphysikproblemen, basierend auf der Methode der finiten Elemente. Neben anderer Software für Computer-Aided Engineering wie ANSYS und ABAQUS wird COMSOL Multiphysics, dessen Anwendung eine immer größere Verbreitung findet, seit über zehn Jahren in vielen Instituten eingesetzt. Daher wurde 2019 am KIT ein dreijähriger "COMSOL Multiphysics Campus-Vertrag" mit einer begrenzten Anzahl an akademischen Forschungslizenzen abgeschlossen. Eine Vertragsform dieser Art bietet die Firma COMSOL Multiphysics GmbH erstmalig an einer deutschen Universität an. Sie ermöglicht ein maßgeschneidertes Lizenzierungsmodell für das KIT und bringt viele zusätzliche Vorteile für die Institute mit sich.

Ngan Long Tu

Die Anwendung COMSOL Multiphysics, ehemals FEMLAB, gibt es in der aktuellen Version 5.5 für die Betriebssysteme Linux, Windows und macOS. Die Software ist modular aufgebaut; d.h. neben dem Basisprodukt „COMSOL Multiphysics“ werden noch mehr als 30 weitere COMSOL Add-On-Module angeboten, die die Simulation einzelner physikalischer Disziplinen ermöglichen, wie z.B. in den Bereichen Elektromagnetik, Strukturmechanik, Akustik, Strömung, Wärmetransport, Verfahrenstechnik u.v.m. (s. Abbildung 1). COMSOL Multiphysics zeichnet sich durch ein modernes, intuitives und einheitliches Bedienungskonzept und die Möglichkeit, gekoppelte Gleichungssysteme verschiedener Disziplinen simultan lösen zu können, aus.

### Akademische COMSOL Multiphysics Lizenzen am KIT

Neben den COMSOL Multiphysics Forschungslizenzen (FNL) stehen ebenso Klassenraumlizenzen (CKL) für die Lehre und für die Poolraumausstattung am SCC zur Verfügung. Im August

2019 konnte das SCC mit der Fa. COMSOL Multiphysics GmbH eine Campus-Lösung für akademische Forschungslizenzen (FNL) am KIT ausarbeiten und dabei alle vorhandenen Einzelverträge in den Instituten am KIT zusammenfassen. Die Voraussetzung dafür war eine enge Zusammenarbeit des SCC mit 16 weiteren Instituten, die bereits COMSOL Multiphysics Lizenzen besitzen und mit dem SCC die Kosten der Campus-Lösung während der Vertragslaufzeit von drei Jahren tragen. Bezogen auf alle am KIT vorhandenen COMSOL-Verträge in 2019 wurde der Leistungsumfang im neuen Campus-Vertrag bei nahezu gleichen Kosten erheblich gesteigert:

- Aufstockung der Hauptlizenzen um 14% (von 66 auf 75) und der Add-On-Lizenzen um ca. 100% (von 220 auf 450)
- zweimal jährlich am KIT stattfindende Schulungskurse in COMSOL Multiphysics
- Supportleistung durch den Vor-Ort-Service erweitert

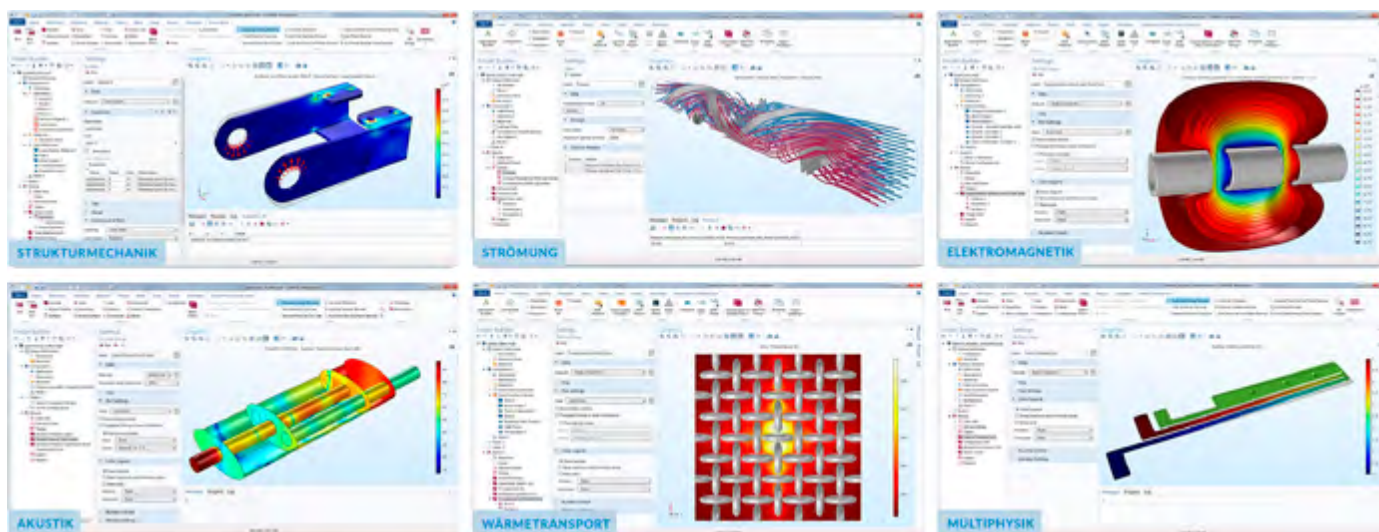


Abbildung 1: COMSOL Multiphysics Add-On-Module, Quelle: [www.comsol.de/comsol-multiphysics](http://www.comsol.de/comsol-multiphysics)

Die ausreichend hohe Anzahl an Forschungslizenzen erhöht zum einen die Effektivität der Simulationsarbeit durch mögliche parallele Abläufe auf unterschiedlichen Rechnersystemen – z. B. auf lokalen Rechnern in den Instituten und zeitgleich auf zentralen HPC-Systemen am SCC.

Zum anderen erlaubt sie die Verwendung des kompletten Portfolios der COMSOL Multiphysics Add-On-Software, wodurch eine hohe Flexibilität im Einsatz der Software für eine große Bandbreite an Forschungsprojekten in den unterschiedlichsten Disziplinen gewährleistet wird (s. Abbildung 2).

Das SCC betreibt Lizenzserver für unterschiedliche Software und stellt auch die akademische Lizenz von COMSOL Multiphysics als Netzwerklizenz für das KIT bereit. Somit übernimmt das

SCC die Aufgaben des Lizenzmanagements, ergänzt durch die Bereitstellung der Installationsmedien auf FTP-Servern und der Installationspakete auf den HPC-Systemen des SCC sowie die fachliche Beratung zur Anwendung. Der Leistungsumfang des Campus-Vertrages enthielt kostenfreie Einführungskurse, die im Oktober 2019 vor Ort und im April 2020 aufgrund der Corona-Virus-Situation online durchgeführt wurden. Diese wurden von zahlreichen Beschäftigten und Studierenden des KIT gerne angenommen. Da das Sommersemester 2020 am KIT von Online-Lehrveranstaltungen geprägt sein wird, können einige Institute auf COMSOL Multiphysics Klassenraumlizenzen zurückgreifen und so einen Teil der praktischen Versuche mittels COMSOL-Simulationen durchführen. Die Studierenden können diese virtuell an ihren eigenen Rechnern nachvollziehen und wiederholen.

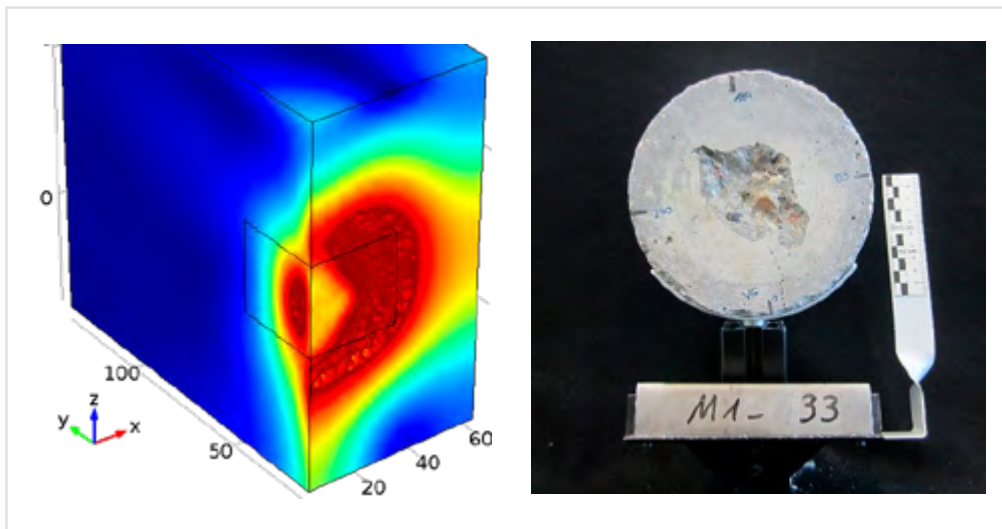


Abbildung 2: „A drying and thermoelastic model for fast microwave heating of concrete“, Quelle: Lepers\_Umminger (gekoppelte Simulation mit COMSOL Multiphysics (Bild links); Versuche am Institut für Hochleistungs- und Mikrowellentechnik (Bild rechts))

#### Weitere Informationen

[1] Webseite zu COMSOL Multiphysics am SCC:

[www.scc.kit.edu/produkte/3850](http://www.scc.kit.edu/produkte/3850)

[2] COMSOL Multiphysics Webseiten:

[www.comsol.de/comsol-multiphysics](http://www.comsol.de/comsol-multiphysics)

[www.comsol.de/products](http://www.comsol.de/products)

#### COMSOL Multiphysics academic campus license at the KIT

In many institutes of the KIT the simulation software COMSOL Multiphysics has been used for over 10 years, and it is becoming more and more widespread. For this reason, a three-year "COMSOL Multiphysics Campus Contract" with a limited number of academic research licenses was concluded at KIT in 2019. This is the first time the company COMSOL Multiphysics GmbH offers a contract of this type at a German university or research institution. This enables a customized licensing model for the KIT and brings many additional advantages for the institutes.

## Business Process Automation mit MS-Orchestrator

Arbeits-, Projekt- und Forschungsgruppen, die sich – oft auch nur für eine bestimmte Zeit – ad-hoc zusammenfinden, benötigen schnell und zuverlässig Plattformen und Infrastrukturen, um effizient zusammenzuarbeiten. Dies erzeugt im IT-Betrieb einen hohen Administrations- und Dokumentationsaufwand bei der Erstellung und Abwicklung von benötigten IT-Ressourcen. Dieser Artikel zeigt am Beispiel der Groupware-Lösung SharePoint, wie durch die Automatisierung von technischen Abläufen, die Business Process Automation, die Effizienz und Transparenz in einer komplexen IT-Umgebung erhöht und gleichzeitig eine gleichbleibende Service- und Dokumentations-Qualität erreicht werden kann.

*Bernd Zakel, Tobias Zuber*

Das SCC stellt als Groupware-Lösung Microsoft SharePoint zur Verfügung. Die gängigen SharePoint-Features gemeinsame Dateiablage/-bearbeitung, Kalender und Listen werden sowohl KIT-intern als auch mit externen Partnern verwendet. Im Laufe der Zeit ist die Anzahl der SharePoint-Teamseiten, die sog. SiteCollections, am KIT stetig gestiegen und damit auch der damit verbundene Administrationsaufwand für die Standard-Prozesse Anlegen, Ändern, Löschen sowie die

anschließende Dokumentation der Seiten. Parallel dazu musste auch die benötigte Systemumgebung ausgebaut werden, um sowohl das Nutzeraufkommen als auch eine möglichst durchgängige Verfügbarkeit des Dienstes sicherstellen zu können.

Es stellten sich daher die Fragen, wie sich wiederkehrende Aufgaben automatisieren lassen, die Dokumentation stets aktuell halten lässt und wie hierbei am SCC verwendete Systeme, Prozesse

und Tools bestmöglich integriert werden können. Zudem sollten sich die Abläufe in die vorhandene IT Service Management-Landschaft einfügen.

### Microsoft System Center Orchestrator

Zur Automatisierung der oben aufgeführten Prozesse kommt die Software System Center Orchestrator von Microsoft (kurz Orchestrator) zum Einsatz. Diese bringt eine Vielzahl an sogenannten Activi-

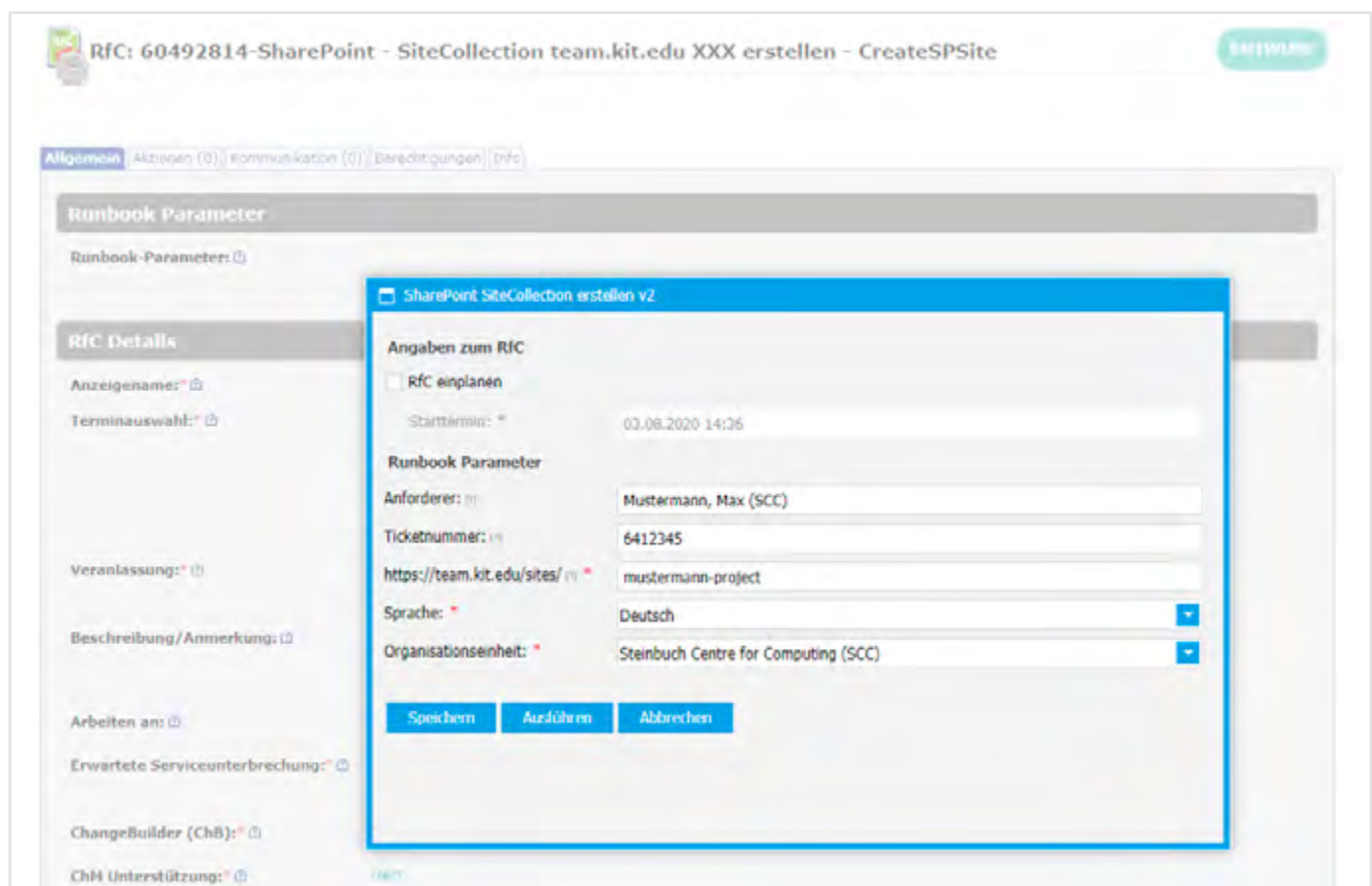


Abbildung 1: RFC zur Erstellung einer SiteCollection mit Parametereingabe

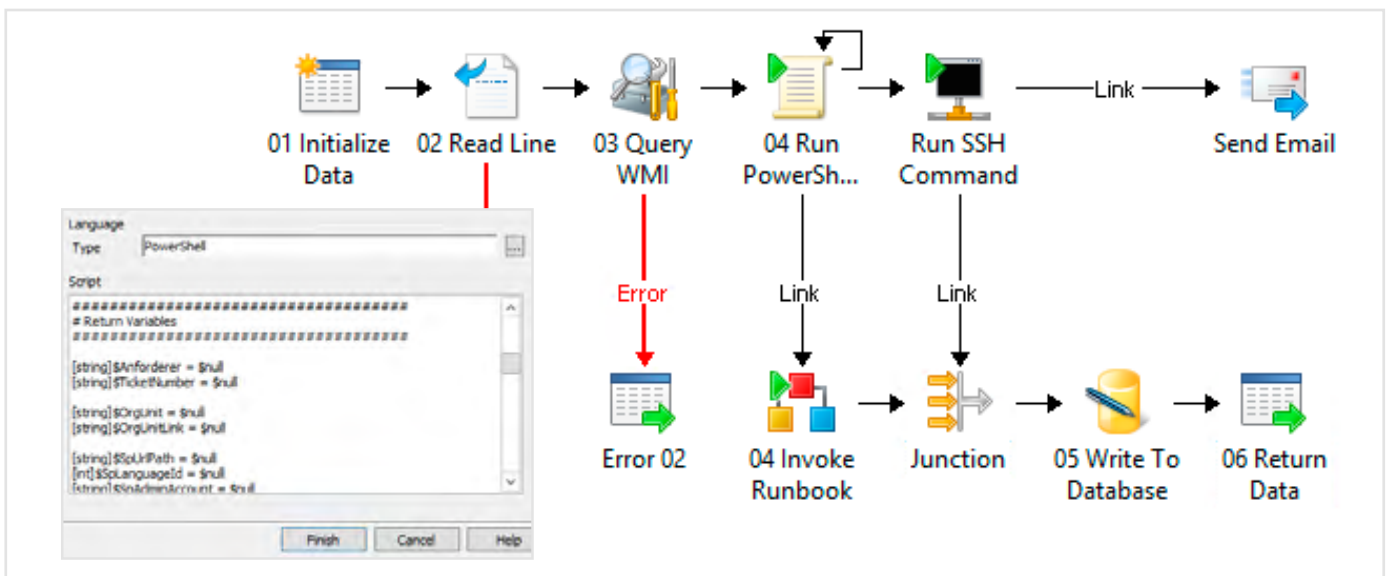


Abbildung 2: Beispiel eines Orchestrator Runbooks im grafischen Runbook Designer

ties mit, um einzelne Funktionen und Aufgaben abbilden zu können (Beispiele für Activities sind u.a. Dateiaktionen, Zugriff auf Webservice-Schnittstellen, Datenbankaktionen, aber auch PowerShell- und Konsolen-Skripte sowie SSH-Zugriffe). Diese Activities können dann in sogenannten Runbooks zu komplexeren Aktionen verbunden und wiederverwendet werden. Zum Erstellen der Runbooks wird ein grafischer Designer mitgeliefert (Abbildung 2). Der Orchestrator bietet die Möglichkeit, weitere Funktionalitäten durch das Importieren sogenannter Integration Packs (IP) hinzuzufügen. IPs können entweder selbst entwickelt oder von Drittanbietern erworben werden.

Damit stehen Prozessdesignern eine große Anzahl an Schnittstellen und Aktionen zur Verfügung, um die unterschiedlichsten IT-Systeme anbinden und steuern zu können.

### IT Service Management am SCC

Eine am SCC entwickelte und an ITIL<sup>1</sup> ausgerichtete webbasierte Service Management-Lösung, die SCC-CMDB, bildet die verschiedenen Management-Prozesse ab. Im Servicekatalog werden IT-Services beschrieben, veröffentlicht und

mit unterschiedlichsten Configuration Items (CI), wie beispielsweise Serversystemen oder IT-Ressourcen (z. B. Datenbanken, SiteCollections, Dokumentationen) in Beziehung gesetzt. Mit Hilfe des Change Management-Verfahrens werden Änderungen am IT-Service oder dessen Komponenten über sogenannte Requests for Change (RfC) registriert, geprüft, freigegeben und rechtzeitig als Wartungsmeldung veröffentlicht. Nach Beendigung der Arbeiten sind die Änderungen am IT-Service und an den Komponenten nachvollziehbar dokumentiert.

### RfC-Templates zur Automatisierung von Standardaufgaben

Zur Automatisierung von Standardaufgaben werden RfC-Templates (Vorlagen) definiert und autorisierten Personen oder IT-Teams zugeordnet.

Soll beispielsweise eine neue SharePoint SiteCollection hinzugefügt werden, erstellen die berechtigten Personen einen neuen RfC basierend auf der entsprechenden Vorlage. Dabei wird ein Eingabefenster zur Erfassung zusätzlicher Parameter und zur Festlegung des Startzeitpunkts geöffnet (Abbildung 1).

Beim Speichern erfolgt ein Eintrag in eine interne Datenbanktabelle, und die Steuerung wird an den Orchestrator übergeben. Dieser wertet den Eintrag aus, startet das Runbook und arbeitet die definierten Activities der Reihe nach ab.

Für das Hinzufügen einer SiteCollection bedeutet dies: das Prüfen der Eingabeparameter, das Erstellen der SiteCollection auf der SharePoint-Farm und einer neuen IT-Ressource in der SCC-CMDB (Abbildung 3) sowie das Verknüpfen der Ressource mit der zugehörigen Datenbank.

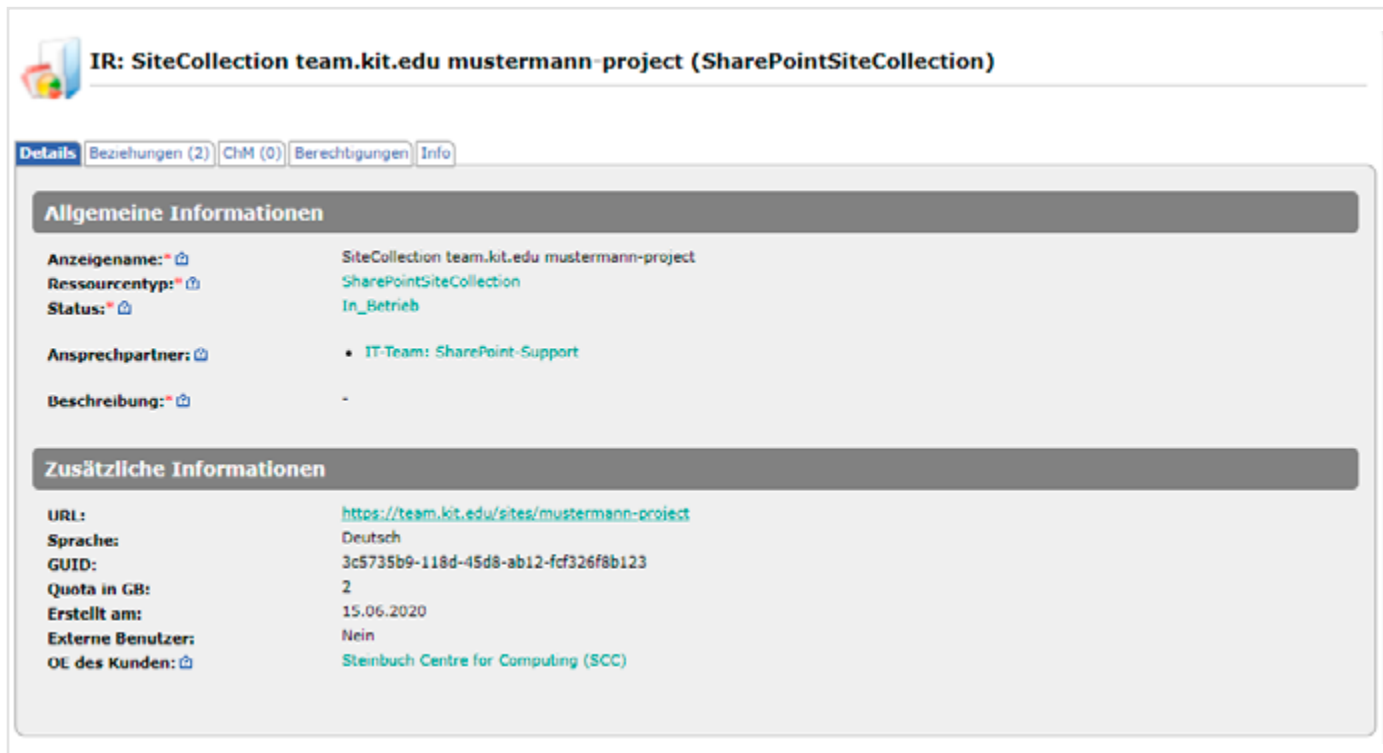
Im Erfolgsfall wird der Status des RfC automatisch auf Durchgeführt gesetzt, andernfalls auf Zurückgewiesen und die durchführende Person (Change Builder) zusätzlich per E-Mail informiert.

### Automatisierte Administration der SharePoint-Farm

Um die Server der SharePoint-Umgebung aktuell zu halten, gibt es RfC-Vorlagen und zugehörige Runbooks. Über diese lassen sich z.B. die Installation von Software oder (Sicherheits-) Updates einplanen und steuern. Die Aktionen werden, wo es möglich ist, nacheinander auf allen Servern durchgeführt, um eine hohe

<sup>1</sup> Die IT Infrastructure Library (ITIL) ist eine international breit genutzte Best-Practice-Sammlung für das IT Service Management und gilt innerhalb der Community als de-facto Standard.





**IR: SiteCollection team.kit.edu mustermann-project (SharePointSiteCollection)**

Details | Beziehungen (2) | ChM (0) | Berechtigungen | Info

**Allgemeine Informationen**

Anzeigename: SiteCollection team.kit.edu mustermann-project  
 Ressourcentyp: SharePointSiteCollection  
 Status: In\_Betrieb  
 Ansprechpartner: IT-Team: SharePoint-Support  
 Beschreibung: -

**Zusätzliche Informationen**

URL: <https://team.kit.edu/sites/mustermann-project>  
 Sprache: Deutsch  
 GUID: 3c5735b9-118d-45d8-ab12-fc326f8b123  
 Quota in GB: 2  
 Erstellt am: 15.06.2020  
 Externe Benutzer: Nein  
 OE des Kunden: Steinbuch Centre for Computing (SCC)

Abbildung 3: Neue SiteCollection in der SCC-CMDB

Verfügbarkeit zu erreichen und Einschränkungen für Benutzer zu minimieren.

Die Aktionen beinhalten u.a. das Setzen bzw. Rücksetzen des Wartungsmodus im System Center Operations Manager. Dies verhindert Alarmierungen für laufende Wartungsarbeiten. Des Weiteren die Deaktivierung bzw. Aktivierung der einzelnen Server im Load Balancer, um Nutzeranfragen auf die noch aktiven Systeme zu leiten. Die eigentliche Installation der Software-Pakete übernimmt der Microsoft Endpoint Configuration Manager – ehemals System Center Configuration Manager. Dieser lässt sich sowohl mit PowerShell-Skripten als auch mit Orchestrator Integration Packs steuern.

### Business Process Automation with MS-Orchestrator

Working, project and research groups need platforms and infrastructures quickly and reliably in order to work together efficiently. In IT operations, this generates a high level of administration and documentation effort when creating and handling the required IT resources. This article takes the groupware solution SharePoint as an example to show how the automation of technical processes, business process automation, can increase efficiency and transparency in a complex IT environment and at the same time achieve a consistent quality of service and documentation.

Um die Gesamtverfügbarkeit der Farm zu gewährleisten, wird jeder Schritt der Wartungsarbeiten protokolliert und von den nachfolgenden Aktionen ausgewertet. Sollte ein Problem festgestellt werden, wird der Wartungsprozess beendet und die Administratoren benachrichtigt. Dadurch bleiben selbst im Fehlerfall ein Teil der Server, und damit der Dienst als Ganzes, funktional. Die Wartungen können somit auch problemlos nachts oder am Wochenende automatisiert ausgeführt werden.

### Zusammenfassung und Ausblick

Durch die Nutzung vorhandener Schnittstellen in Verbindung mit PowerShell

ist es möglich, einige für den Betrieb der SharePoint-Umgebung notwendige Abläufe zu automatisieren und in die am SCC bestehende IT Service Management-Landschaft zu integrieren. Microsoft System Center Orchestrator stellt hierfür eine große Auswahl von Activities und Erweiterungsmöglichkeiten bereit. Somit konnte eine robuste und fehlertolerante Prozessautomatisierung für häufig wiederkehrende Aufgaben mit wiederholbaren Prozess-Ergebnissen umgesetzt werden. Ein weiterer Vorteil der Lösung ist, dass sich damit auch der manuelle Aufwand für die Dokumentation drastisch verringert hat.

Aktuell wird evaluiert, wie die beschriebenen Softwarekomponenten zur Prozessautomatisierung für die Verwaltung des gesamten Lebenszyklus von Endgeräten (Client Management) eingesetzt werden kann.

# Upgrade auf Microsoft Exchange 2019

Das SCC hat Anfang dieses Jahres die zentralen Postfachserver erfolgreich von Microsoft Exchange Server 2013 auf Microsoft Exchange Server 2019 aktualisiert. Damit erhalten alle Beschäftigten und Studierenden am KIT eine aktuelle und in der Sicherheit verbesserte Kommunikationsplattform sowie mehr Speicherplatz für ihre Postfächer.

*Eugenie Bantle, Michael Willhauk*

## Einleitung



Eines der am häufigsten genutzten Kommunikationsmittel am KIT ist der Austausch von E-Mails. Das SCC betreibt bereits seit vielen Jahren dafür als Groupware-Lösung Microsoft Exchange, womit weitere Funktionalitäten wie Termin- und Kontaktverwaltung bereitgestellt werden.

Um diesen essentiellen Dienst aktuell zu halten, wurden die Microsoft Exchange Server auf die Version 2019 aktualisiert und dabei auch die dafür notwendige Hardware erneuert.

## Vorbereitungen der Migration

Die Aktualisierung des zentralen E-Mail-Dienstes bedarf einer sorgfältigen Planungs- und Testphase. Dazu wurde Anfang 2019 die bestehende Exchange-Testumgebung mit neuen Exchange 2019 Servern erweitert. Zur Installation und Konfiguration der neuen Exchange 2019 Testserver wurden PowerShell-Skripte entwickelt, die später bei der Migration der Produktionsumgebung wiederverwendet werden konnten. Wie von Microsoft empfohlen, wird für die neuen Server als Betriebssystem Microsoft Windows 2019 Server Core eingesetzt. Mit Microsoft Server Core wird die Sicherheit durch die reduzierte Angriffsfläche erhöht und der Wartungsaufwand für die einzelnen Systeme verringert.

Das bei Exchange 2013 eingesetzte Redundanzkonzept mit aktiven und passiven Datenbankkopien, die auf die beiden Standorte Campus Nord und Campus

Campus Nord			Campus Süd		
Exchangeserver	Aktive DB	Passive DB	Exchangeserver	Aktive DB	Passive DB
KIT-MSX-41	DB4142A	DB4241A	KIT-MSX-42	DB4241A	DB4142A
	DB4142B	DB4241B		DB4241B	DB4142B
	...	...		...	...
	DB4142J	DB4241J		DB4241J	DB4142J
KIT-MSX-43	DB4344A	DB4443A	KIT-MSX-44	DB4443A	DB4344A
	DB4344B	DB4443B		DB4443B	DB4344B
	...	...		...	...
	DB4344J	DB4443J		DB4443J	DB4344J
...			...		

Abbildung 1: Schematische Darstellung der Datenbankverteilung

Süd aufgeteilt sind (Abbildung 1), sowie die Virtualisierung der Exchange Server mit VMware ESXi hat sich bewährt und wurde auch für die Migration auf Exchange 2019 beibehalten.

Für die Produktionsumgebung wurden in der Virtualisierungsumgebung des SCC sechs Hosts beschafft, auf denen 12 neue Exchange 2019 Server installiert werden sollten.

Spezifikation eines ESXi-Hosts:  
 2 x Intel Xeon Gold 5120  
 768 GB RAM  
 2 x 25 GE Ethernet  
 2 x 16 GBit/s FibreChannel

Auf einem Host laufen im Normalzustand zwei virtuelle Maschinen (VMs) mit Exchange 2019 Server. Die Kapazität der Hosts wurde so dimensioniert, dass bei Ausfall eines Hosts (z.B. bei Updates der Virtualisierungssoftware ESXi) dessen VMs auf den verbliebenen zwei Hosts im selben Standort weiterlaufen können (Abbildung 2).

## Durchführung der Migration

Ende 2019 standen die virtuellen Maschinen auf der neuen Plattform bereit. In der Weihnachtspause installierte das Exchange-Team zügig die neuen

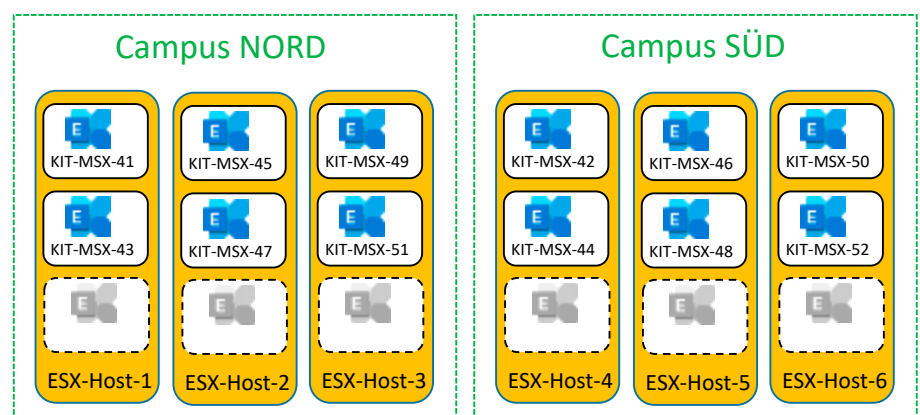


Abbildung 2: Aufteilung der virtuellen Maschinen

Exchange Server mit Hilfe der zuvor in der Testumgebung erstellten Skripte, konfigurierte sie und nahm sie in die vorhandene Exchange-Organisation auf.

Ein kritischer Punkt bei der Migration war die Umstellung des Clientzugriffs Mitte Januar 2020, da sich ab diesem Zeitpunkt alle Clients mit den neuen Exchange Servern verbunden haben. Da der bestehende Namensraum weiterverwendet wurde, mussten die Benutzerinnen und Benutzer keine Änderung an ihrem Mailclient vornehmen.

Nach einer kurzen Pilotphase begann Anfang Februar die Umstellung der Postfächer. Dabei wurden täglich Postfachdaten von ca. 2 Terabyte auf die

neuen Server verschoben. Auch diese Umstellung erfolgte für die Benutzerinnen und Benutzer auswirkungs- und unterbrechungsfrei, da der Postfachinhalt im Hintergrund auf die neuen Server synchronisiert wurde, bevor der endgültige „Switchover“ stattfand.

Insgesamt wurden fast 60.000 Postfächer verschoben, die aktuell ein (aktives) Datenvolumen von ca. 48 Terabyte einnehmen.

### Nur noch TLS 1.2

Die Protokolle TLS 1.1 und TLS 1.0 gelten schon länger als unsicher. Exchange 2019 unterstützt deshalb nur noch TLS 1.2.

Ältere Clients, die diesen Anforderungen nicht genügen, können seither nicht mehr zugreifen. Dies betrifft beispielsweise:

Windows 7 Rechner: Outlook verwendet standardmäßig kein TLS 1.2

- MacOS Rechner älter als Sierra 10.12: Mac Mail unterstützt kein TLS 1.2
- Android älter als 4.2: Mail App unterstützt kein TLS 1.2

### Mehr Platz für die Postfächer

Die Standardquota der Postfächer wurde auf 2 Gigabyte für Studierende und auf 8 Gigabyte für Beschäftigte, Gäste und Partner sowie Funktionspostfächer erhöht.



Abbildung 3: Die Weboberfläche von Exchange 2019 hat sich in der Funktionalität und Usability weiterentwickelt.

Weitere Informationen: [docs.microsoft.com/de-de/exchange/new-features/new-features?view=exchserver-2019#clients](https://docs.microsoft.com/de-de/exchange/new-features/new-features?view=exchserver-2019#clients)

### Upgrade to Microsoft Exchange 2019

At the beginning of this year, the SCC successfully updated the central mailbox servers from Microsoft Exchange Server 2013 to Microsoft Exchange Server 2019. The hardware has been renewed, too.

The redundancy concept used for Exchange 2013, with active and passive database copies distributed between the two locations Campus North and Campus South, as well as virtualization with VMware ESXi of the Exchange servers, has proven itself and has been retained for Exchange 2019.

During the Christmas break, the new Exchange servers were quickly installed, configured and added to the existing Exchange organization using scripts previously created in the test environment. A critical point was the changeover of client access in mid-January 2020, as all users connected to the new Exchange servers from that time on. For the users this change was transparent, because the existing namespace was still used and therefore no changes had to be made on the client side.

After a short pilot phase, the moving of the mailboxes started at the beginning of February. By the end of March, about 60,000 mailboxes were moved, which currently take up an (active) data volume of about 48 Terabyte.

## Effiziente Vorgangsbearbeitung mit dem SCC-Ticketsystem

Das SCC bietet seit dem 15.05.2020 jeder Organisationseinheit und deren Support-Teams am KIT an, das SCC-Ticketsystem zur Unterstützung der eigenen Support-Prozesse einzusetzen und damit eine effiziente und komfortable Methode der Vorgangsbearbeitung zu nutzen.

Birgit Junker

Zur Erfassung und Bearbeitung von Nutzeranfragen setzt das SCC bereits seit 2011 ein Ticketsystem ein. Dieses wurde speziell für die Vorgangsbearbeitung im IT-Support entwickelt und im Jahr 2018 auf das im IT-Support vielfach als Helpdesk-System eingesetzte Service-Management-System OTRS umgestellt (s. SCC-News 2/2018, S. 16).

Zu Beginn konnten in diesem System nur Beschäftigte des SCC und am KIT benannte IT-Beauftragte, die IT-Fachleute in den Organisationseinheiten (OE), erfassen. Mit Einführung des neuen, auf OTRS basierenden Ticketsystems wurden schrittweise weitere Nutzerkreise berechtigt.

Mittlerweile haben alle Angehörigen des KIT (Beschäftigte, IT-Beauftragte, Studierende, Gäste und Partner) mit einem gültigen KIT-Account einen Zugang und damit die Möglichkeit, Anfragen direkt über die webbasierte Nutzerschnittstelle an das SCC zu stellen.

An den Service Desk adressierte Anrufe oder E-Mails, die dort nicht direkt beantwortet oder gelöst werden können, werden im System als Vorgang (Ticket) erfasst und bei Bedarf zur weiteren Bearbeitung an die Fachabteilungen weitergegeben. Die Weitergabe erfolgt über sogenannte Ticket-Queues (Warteschlangen für Vorgänge), die themenspezifisch angelegt sind (s. Abbildung 1).

Nachdem ein neuer Vorgang gesichtet und einer themenspezifischen Ticket-Queue zugeordnet wurde, erhalten die Personen des zuständigen Support-Teams eine Mail mit Informationen zum Vorgang.

Ein Team-Mitglied übernimmt die Bearbeitung des Vorgangs und kommuniziert über die webbasierte OTRS-Agentenschnittstelle mit der anfragenden Person oder mit Kolleginnen und Kollegen. Dort können auch telefonische Absprachen dokumentiert, Notizen hinzugefügt oder

der Vorgang für weitere Tätigkeiten an ein anderes Support-Team weitergegeben werden.

So bietet das SCC-Ticketsystem allen am Supportprozess beteiligten Personen eine Übersicht über selbst eingestellte Anfragen oder zu bearbeitende und erledigte Vorgänge. Die aus dem Identitätsmanagementsystem (IDM) stammenden Benutzerinformationen geben dem Support-Team Auskunft über die OE-Zugehörigkeit des Benutzers und seine offenen Vorgänge. Der Zugriff auf die Vorgänge wird über ein Berechtigungskonzept auf Team-Ebene definiert. Für jede Queue wird festgelegt, welches Support-Team darin Vorgänge neu erstellen, verschieben, lesen oder bearbeiten darf.

### Voraussetzungen für den Einsatz

Für den Einsatz des SCC-Ticketsystems in OEs wurde eine Dienstvereinbarung am KIT abgeschlossen. Darin sind die Einführung und Anwendung eines einheitlichen Ticketsystems in der Dienststelle geregelt. Definiert werden u.a. die Zweckbestimmung des Ticketsystems, Beschreibung des Verfahrens und der Zugriffsberechtigungen, die Datensicherung/-speicherung sowie die Unzulässigkeit einer möglichen Leistungs- und Verhaltenskontrolle im Rahmen des Beschäftigungsverhältnisses. Desweiteren muss die verantwortliche OE Informationen zur Nutzung im elektronischen Verzeichnis (eVV) des Datenschutzteams dokumentieren (s. SCC-News 1/2018 S. 17). Das SCC stellt dazu eine Vorlage bereit.

Vorgänge im SCC-Ticketsystem dürfen ausschließlich von Angehörigen des KIT bearbeitet werden, während Support-Anfragen auch von Externen eingestellt werden können. Dazu richtet das SCC in

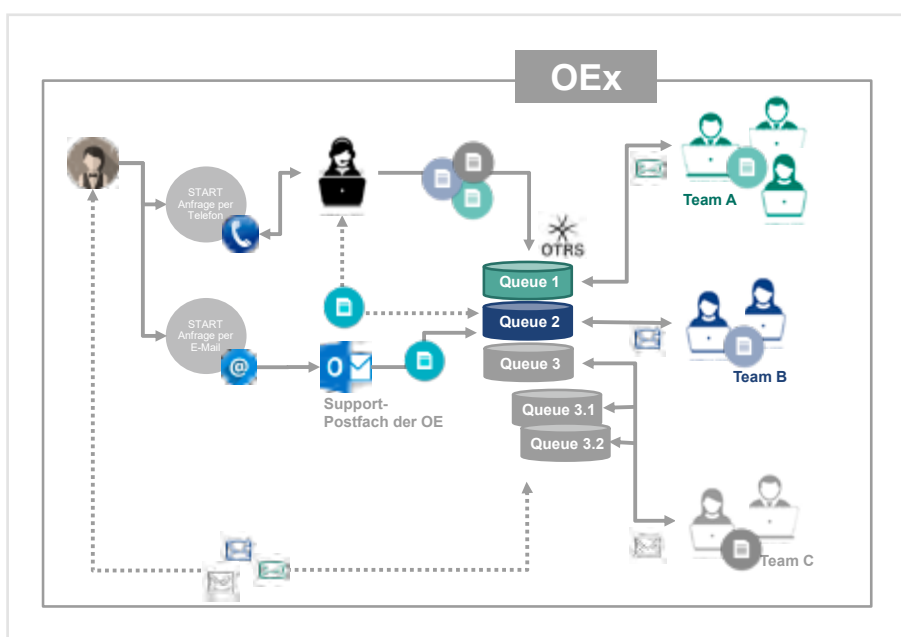


Abbildung 1: Unterstützung der Supportprozesse einer KIT-OE mit dem SCC-Ticketsystem



Abgabe mit der OE ein Support-Postfach ein. Eine an dieses Postfach gesendete Mail wird automatisch zu einem Vorgang im Ticketsystem.

Zudem sind die Team-spezifischen Supportprozesse sowie deren Daten für andere Support-Teams nicht einsehbar (Abbildung 2). Die Daten und Prozesse werden zwar von ein und demselben System gesteuert und nutzen die gleichen Nutzerdaten, diese können aber nicht gegenseitig eingesehen werden. Das Berechtigungskonzept erlaubt jedoch das

Einrichten OE-übergreifender Teams, die Anfragen in einer Ticket-Queue gemeinsam bearbeiten können.

#### Unterstützung durch das SCC

Das SCC unterstützt die OEs bei der Umsetzung der OE-spezifischen Anforderungen sowie beim Erstellen der Datenschutz-relevanten Dokumentation. Zudem klärt das SCC in Rücksprache mit der OE die Team-spezifischen Konfigurationen und setzt diese entsprechend im System um. Mitarbeiterinnen und Mitarbeiter von

Vorgängen werden in die OTRS-Agentenschnittstelle eingewiesen. Das erleichtert den OEs und Support-Teams den Einstieg in das neue Arbeitsmittel.

Damit bietet das SCC am KIT allen Organisationseinheiten die Möglichkeit, ein Ticketsystem für die OE-eigene Vorgangsbearbeitung einzusetzen, das, wie es in der Dienstvereinbarung heißt, „zur Förderung einer effizienten, kooperativen, vertrauensvollen und nutzerorientierten Erbringung der Dienstleistungen“ beiträgt.

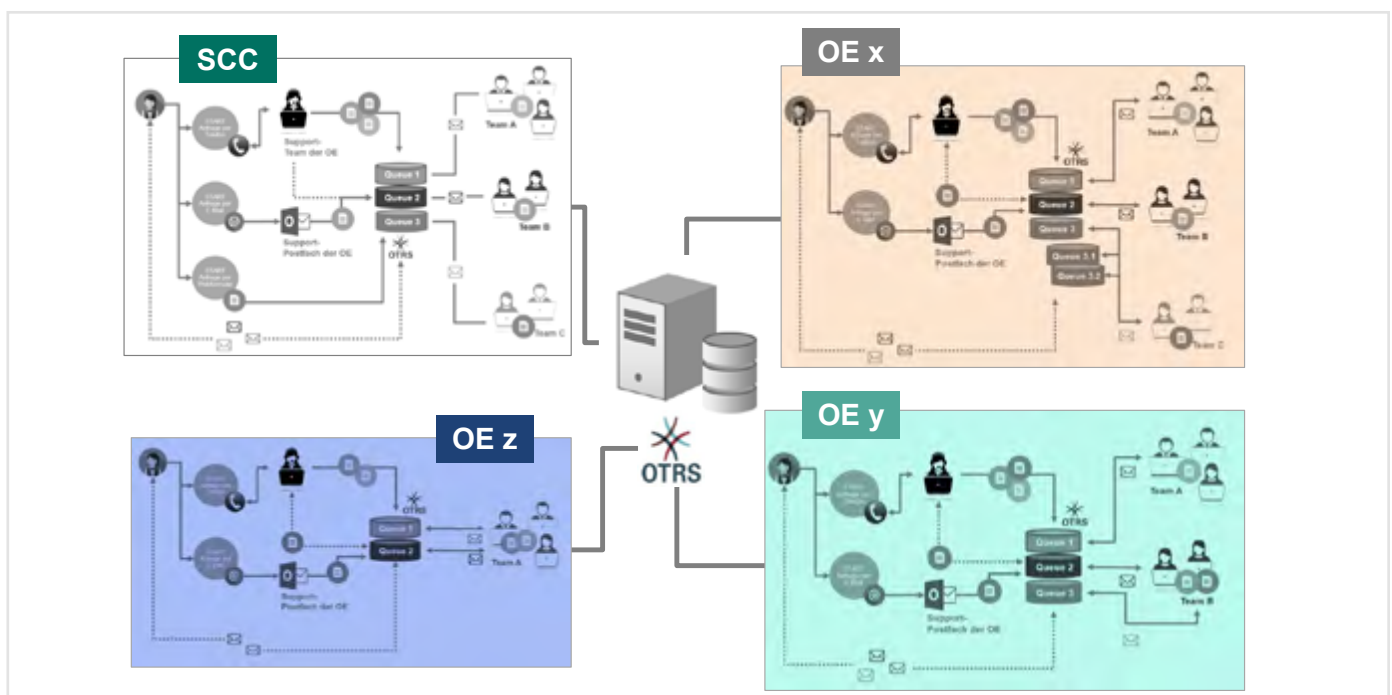


Abbildung 2: Trennung der Supportprozesse unterschiedlicher Support-Teams

#### Efficient support with the SCC ticket system

In the meantime, all KIT members (employees, IT appointees, students, guests and partners) with a valid KIT account have access and the possibility to submit inquiries directly to the SCC via the customer login.

Once a case has been categorized and assigned to a ticket queue, the members of the responsible IT team receive a mail with information and a reference to the case.

For processing, they use the agent interface, via which they communicate with the requestor by mail or manually document telephone contacts.

The visibility of the processes is defined using the authorization concept at team level. For each queue, you define which IT team can create, move, and read or edit new cases in it.

Since May 15, 2020, SCC offers each organizational unit (OU) and their support teams at KIT to use the SCC ticket system to manage their own support processes and to use this efficient method of support case processing.

#### Requirements for use?

Service agreement defined at KIT since March 2020 contains purpose, description of procedures and roles and regulates the performance and behavior control for all support units. Requests can only be processed by members of KIT, while support requests can also be submitted by non-KIT members.

#### How does SCC support?

SCC checks the team-specific requirements for feasibility, supports creating the documentation relevant to data protection, clarifies the team-specific configurations in consultation with the OU, implements them and introduces the employees.

# HoreKa!

Das SCC beschafft ein neues Tier-2 HPC-System. Der neue „Hochleistungsrechner Karlsruhe“ (kurz HoreKa) wird 2021 voraussichtlich zu den zehn leistungsfähigsten Rechnern Europas gehören und eine Rechenleistung von ca. 17 PetaFLOPS erbringen.

*Simon Raffener*

Von der Energiewende zu neuen Materialien, von der Astrophysik zu den Lebenswissenschaften: Um natürliche und technische Vorgänge in ihrer ganzen Komplexität zu verstehen, nutzen Forscherinnen und Forscher die schnellsten Hochleistungsrechner der Welt. Ab Januar 2021 nimmt das SCC den neuen Supercomputer schrittweise in Betrieb. Das Gesamtsystem wird im Sommer 2021 der Wissenschaft übergeben. Der im Mai unterzeichnete Liefervertrag hat eine Größenordnung von 15 Millionen Euro. Das System wird Wissenschaftlerinnen und Wissenschaftlern aus ganz Deutschland zur Verfügung stehen. Vor allem in den Materialwissenschaften, den Erdsystemwissenschaften, der Energie- und Mobilitätsforschung im Ingenieurwesen, den Lebenswissenschaften sowie der Teilchen- und Astroteilchenphysik werden Forschende dank des neuen Supercomputers ein detaillierteres Verständnis hochkomplexer natürlicher und technischer Vorgänge erlangen können. Selbstverständlich kann HoreKa bei Bedarf auch von Wissenschaftlern genutzt werden, die zum Verständnis des Virus SARS-CoV-2 forschen und damit zur Bekämpfung der Krankheit COVID-19 beitragen.

Das Ausschreibungsverfahren war bewusst technologieoffen gehalten, um das Know-how der Bieter einzufordern und das leistungsfähigste Gesamtsystem zu erhalten. Herausgekommen ist ein innovatives Hybrid-System mit fast 60.000 Intel Xeon Scalable Prozessorkernen der nächsten Generation und mehr als 240 Terabyte Hauptspeicher sowie 740 NVIDIA A100 Tensor Core

GPUs der nächsten Generation. Als Kommunikationsnetzwerk kommt ein non-blocking NVIDIA Mellanox InfiniBand-HDR-Netzwerk mit 200 GBit/s pro Port zum Einsatz.

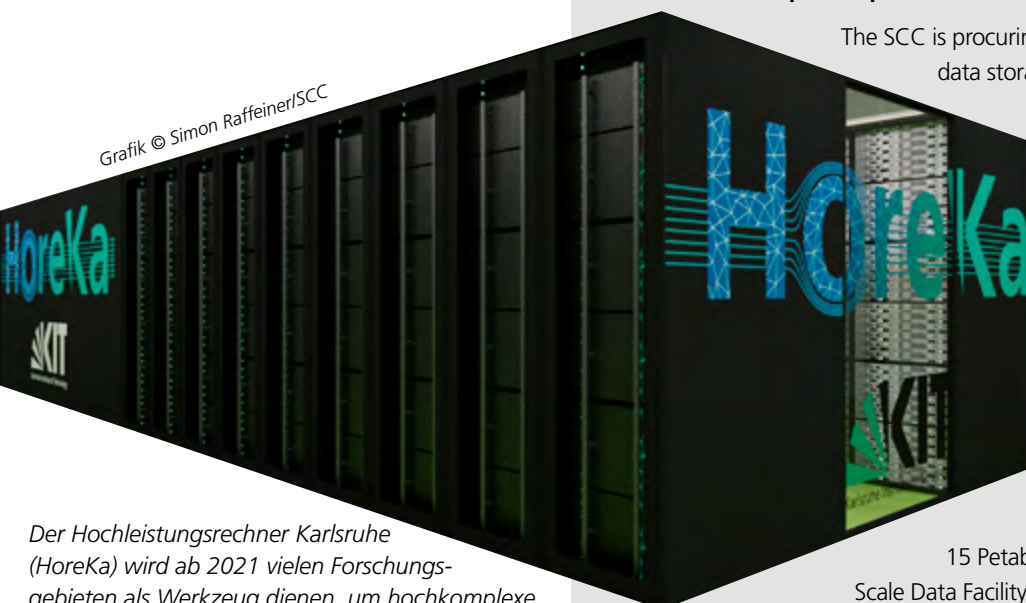
Ein zentraler Gesichtspunkt bei der Auslegung des Systems waren auch die enormen Datenmengen, welche bei wissenschaftlichen Forschungsprojekten anfallen. Je nach Anwendung können von einer einzigen Simulation mehrere Hundert Terabyte an Daten erzeugt werden. Um mit den wachsenden Datenmengen Schritt zu halten, liefern die Rechenknoten, das InfiniBand-Netzwerk und die parallelen Dateisysteme von HoreKa im Vergleich zum Vorgängersystem ForHLR jeweils einen bis zu vier Mal höheren Speicherdurchsatz. Eine mehrstufige Datenhaltung soll zusätzlich die Weiterverarbeitung auf externen Speichersystemen mit hohem Durchsatz garantieren.

Als Datenablage dienen zwei parallele Spectrum-Scale-Dateisysteme mit einer Gesamtkapazität von mehr als 15 Petabyte. HoreKa ist auch mit bis zu 45 GByte/s Datenrate an den LSDF Online Storage angebunden, der seit 2010 eine moderne Infrastruktur für die Speicherung, Verwaltung, Archivierung und Analyse von Forschungsdaten bietet und vom SCC betrieben wird.

HoreKa wird vollständig im 2015 für den Vorgänger ForHLR neu errichteten Rechnergebäude auf dem Campus Nord des KIT untergebracht. Das preisgekrönte, energieeffiziente Heißwasser-Kühlkonzept wird mit dem neuen System fortgeführt.

Den Namen HoreKa wählten die Mitarbeiterinnen und Mitarbeiter des SCC in Anlehnung an „GridKa“, dem „Grid Computing Centre Karlsruhe“. Dieses befindet sich ebenfalls am SCC und stellt seit mehr als 15 Jahren erfolgreich Datenspeicher und Analysekapazitäten für Großexperimente auf der ganzen Welt bereit, auch für den Large Hadron Collider (LHC) am CERN in der Schweiz.

## HoreKa! New Supercomputer at KIT



The SCC is procuring a new Tier 2 HPC cluster system including a new data storage system. The system is called "Hochleistungsrechner Karlsruhe", short „HoreKa“. The first phase will be delivered in autumn of 2020, the full system is expected to be one of the ten most powerful computers in Europe in 2021. The new system will have more than 60.000 Intel Xeon processor cores, more than 240 Terabytes of RAM as well as 740 NVIDIA A100 Tensor Core GPUs and is expected to deliver more than 17 PetaFLOP/s. A non-blocking InfiniBand HDR network with 200 GBit/s per port is used for communication between the nodes. Two Spectrum Scale parallel file systems will offer a total storage capacity of more than

15 Petabytes. HoreKa will also be connected to the "Large Scale Data Facility" (LSDF) with a data rate of up to 45 Gbyte/s. It will

be installed in the state-of-the-art data center constructed for its predecessor ForHLR on KIT's Campus North in 2015.

Der Hochleistungsrechner Karlsruhe (HoreKa) wird ab 2021 vielen Forschungsgebieten als Werkzeug dienen, um hochkomplexe Systeme zu verstehen.



## bwUniCluster 2.0: Neues Tier-3 HPC-System am KIT

Mit dem bwUniCluster 2.0 hat das SCC ein neues Tier-3 HPC-Cluster für die lokale und regionale Grundversorgung mit Rechenleistung in Betrieb genommen. Damit wurde der seit 2014 erfolgreich betriebene bwUniCluster als gemeinsames System aller neun Landesuniversitäten und der im „Hochschulen für Angewandte Wissenschaften e.V.“ organisierten Hochschulen Baden-Württembergs abgelöst.

*Simon Raffener*

Das neue HPC-System firmiert unter dem Namen „bwUniCluster 2.0+GFB-HPC“ und steht, wie das Vorgängersystem, allen Angehörigen der beteiligten Institutionen zur Verfügung. Mit der hohen Rechenleistung von mehr als 2 PetaFLOPS kann aber gleichzeitig auch der Rechenbedarf im Großforschungsbereich des KIT gedeckt werden, zudem wurde der Cluster in die Inf-

rastruktur der Helmholtz Data Federation<sup>1</sup> integriert. Durch die Bündelung der Finanzmittel erhalten alle Beteiligten Zugang zu einem insgesamt leistungsfähigeren System. Gleichzeitig verringert sich für das SCC der Aufwand für die Administration und die Nutzerunterstützung, insbesondere bei den wissenschaftlichen Anwendungen.

► Der neue Cluster ging – trotz der durch die Corona-Pandemie verursachten Einschränkungen – am **17.03.2020** in den planmäßigen Nutzerbetrieb. ◀





Die Neubeschaffung (Abbildung 1) umfasste insgesamt 490 Rechenknoten. Knapp ein Fünftel davon ist für sogenanntes High Throughput Computing ausgelegt, also Rechenjobs, welche nur einzelne Cores oder maximal einen ganzen Knoten belegen. Die übrigen Knoten zielen auf hochparallele Jobs mit Hunderten oder Tausenden von Cores und hohem Kommunikationsbedarf zwischen den einzelnen Knoten. Daher kommt als Verbindungnetzwerk InfiniBand mit einer Geschwindigkeit von 100 Gigabit pro Sekunde zum Einsatz.

Alle Knoten sind mit jeweils zwei Intel-Xeon-Prozessoren der „Cascade Lake“-Generation mit 20 Cores pro Sockel ausgestattet. Der Speicherausbau liegt bei 96 Gigabyte RAM pro Knoten. Für besonders speicherintensive Anwendungen stehen daneben sechs Knoten mit jeweils 3 Terabyte RAM und 4,8 Terabyte NVMe-Speicher zur Verfügung. Weitere 24 Knoten enthalten insgesamt 136 GPU-Beschleuniger des Typs Tesla V100 von NVIDIA. Die NVLink-Technologie erlaubt eine schnelle Kommunikation der Grafikprozessoren untereinander, wodurch

Anwendungen weiter beschleunigt werden können. Die GPU-Knoten decken die stetig steigenden Anforderungen des Machine Learning ab und ermöglichen die Entwicklung neuartiger numerischer Algorithmen.

Die im November 2016 beschaffte Erweiterung des bwUniCluster ist im neuen System aufgegangen. Daher ergibt sich ein Gesamtsystem mit 842 Knoten, mehr als 29.000 CPU-Cores unterschiedlicher Prozessorgenerationen sowie mehr als 108 Terabyte RAM. Das zentrale Datenhaltungssystem hat eine Kapazität größer 5 Petabyte mit einem aggregierten Datendurchsatz von 72 Gigabyte pro Sekunde. Als Softwarelösung kommt wie bei den existierenden Systemen Lustre zum Einsatz. Für KIT-Nutzer existiert eine direkte Anbindung an den LSDF Online Storage, einen Massenspeicher für wissenschaftliche Daten. Zudem steht mit BeeOND erstmals ein lokales, paralleles On-Demand-Dateisystem zur Verfügung. Die mehrstufige Speicherhierarchie ermöglicht es den Nutzern, das für den jeweiligen Anwendungsfall am besten geeignete Speichersystem selbst zu wählen und Optimierungen vorzunehmen.

Die Softwareversorgung und Nutzerunterstützung erfolgen gemeinsam durch die beteiligten Standorte. Die Koordination findet wie beim Vorgänger im Rahmen des vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg geförderten Projektes bwHPC-S5 statt. Auch die bewährten Nutzerunterstützungswege, wie z.B. das bwSupportPortal, werden für das neue System beibehalten.

<sup>1</sup> [www.helmholtz.de/hdf](http://www.helmholtz.de/hdf)

### **bwUniCluster 2.0: New Tier 3 HPC system at KIT**

The SCC has commissioned a new Tier 3 HPC cluster system including a new data storage system for the renewal of the existing bwUniCluster 1. The system is called "bwUniCluster 2.0+GFB-HPC" and has been available to users since March 12, 2020. It is open to members of the nine universities and some of the applied universities in Baden-Württemberg, but also serves KIT's own large-scale research sector (Großforschungsbereich, GFB).

The new cluster consists of 490 nodes, each containing two latest generation "Cascade Lake" Intel processors with 20 cores per CPU and 96 Gigabytes of RAM per node. Six additional nodes with 3 Terabytes of RAM and 4.8 Terabytes of NVMe storage each are available for memory-intensive applications. A further 24 nodes are equipped with NVIDIA Tesla V100 GPU accelerators for machine learning and the development of novel numerical algorithms. The storage system has a capacity of 5 petabytes with a throughput of 72 GB/s. KIT users profit from a direct connection to the Large Scale Data Facility (LSDF).

*Der bwUniCluster 2.0 in den Räumen des SCC am Campus Süd.*





## Anycast für die DNS-Resolver

Für die DNS-Resolver-Adressen, welche auf nahezu allen Systemen für die DNS-Auflösung genutzt werden, ist eine sehr hohe Verfügbarkeit erforderlich. Da eine clientseitige Umschaltung auf die zweite „Backup-Adresse“ nicht immer reibungslos funktioniert, führt die Störung einer Resolver-Adresse mitunter zu großen Einschränkungen. Mit der Einführung der Anycast-Technologie für die Adressen der Resolver wird nun eine sehr hohe Verfügbarkeit sichergestellt.

*Klara Mall, Benedikt Neuffer, Julian Schuh*

### DNS-Auflösung am KIT

Im Normalfall gibt es zwei DNS-Resolver-Adressen, die auf Endgeräten (Clients und Servern) eingetragen werden, wenn diese ihre Konfiguration nicht automatisch via DHCP oder Router Advertisements beziehen. So war es auch an den Vorgängereinrichtungen des KIT, Universität und Forschungszentrum Karlsruhe. Hierbei handelte es sich um die IPv4-Adressen 129.13.64.5 und 129.13.96.2 bzw. 141.52.3.3 und 141.52.8.18. Für IPv6 wurden die Adressen 2a00:1398::2 und 2a00:1398::4 bzw. 2a00:1398::1 und 2a00:1398::3 hinzugenommen. Durch den Zusammenschluss der beiden Einrichtungen gab es also sowohl für IPv4 als auch für IPv6 jeweils vier Resolver-Adressen. Zunächst handelte es sich bei den Resolver-Adressen am KIT immer um die Host-Adressen der zugehörigen Server.

Die hohe Verfügbarkeit der DNS-Resolver über die bekannten IP-Adressen ist kritischer Bestandteil eines zuverlässigen Netzbetriebs, denn jede Wartung oder Störung, die zur Nichtverfügbarkeit dieser Adressen führt, hat große Auswirkungen auf Dienste und Nutzer. Sind die Adressen als Host-Adressen auf den Servern konfiguriert, kommt es zwingend zum Ausfall, sobald ein Server aufgrund einer Wartung oder Störung nicht verfügbar ist.

### Statisches Routing der Resolver-Adressen

Seit 2013 wurden die IP-Adressen der Resolver zur Verbesserung der Verfügbarkeit auf andere Weise implementiert. Die Resolver-Adressen sind nun nicht mehr als Host-Adressen auf den Servern sondern als eigenständige Service-Adressen konfiguriert. Die Host-Adressen beziehen die Server jeweils aus einem dedizierten IPv6- und IPv4-Subnetz. Anfragen an die Service-Adressen werden mit Hilfe von auf den Routern installierten statischen Routen an die Host-Adressen der Maschinen weitergeleitet. Beispielsweise wurden Anfragen an die Adresse 2a00:1398::1 mittels statischer Route an die Host-Ad-

resse 2a00:1398:8:10::8d34:e942 weitergeleitet. Da alle vier Server so konfiguriert waren, dass sie Anfragen an alle Resolver-Adressen beantworten, konnte im Störungs- oder Wartungsfall die entsprechende Route auf einen anderen Server gesetzt werden, so dass Ausfälle behoben bzw. vermieden werden konnten.

### Dynamisches Routing der Resolver-Adressen mit Anycast

Mit der Einführung der Anycast-Technologie für die Resolver-Adressen haben wir kürzlich die Ausfallsicherheit weiter verbessert: Statische Routen, welche bei Ausfall oder Wartung eines Systems manuell angepasst werden müssen, entfallen. Stattdessen kommunizieren

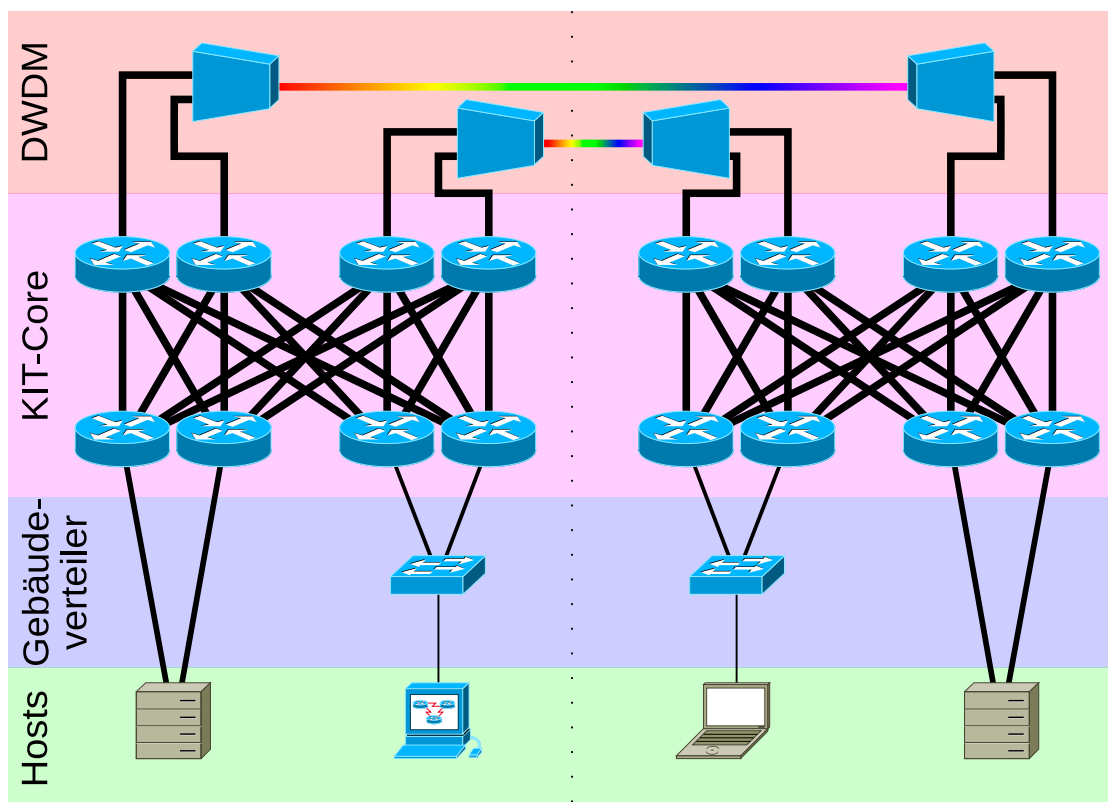


Abbildung 1: Im Normalfall

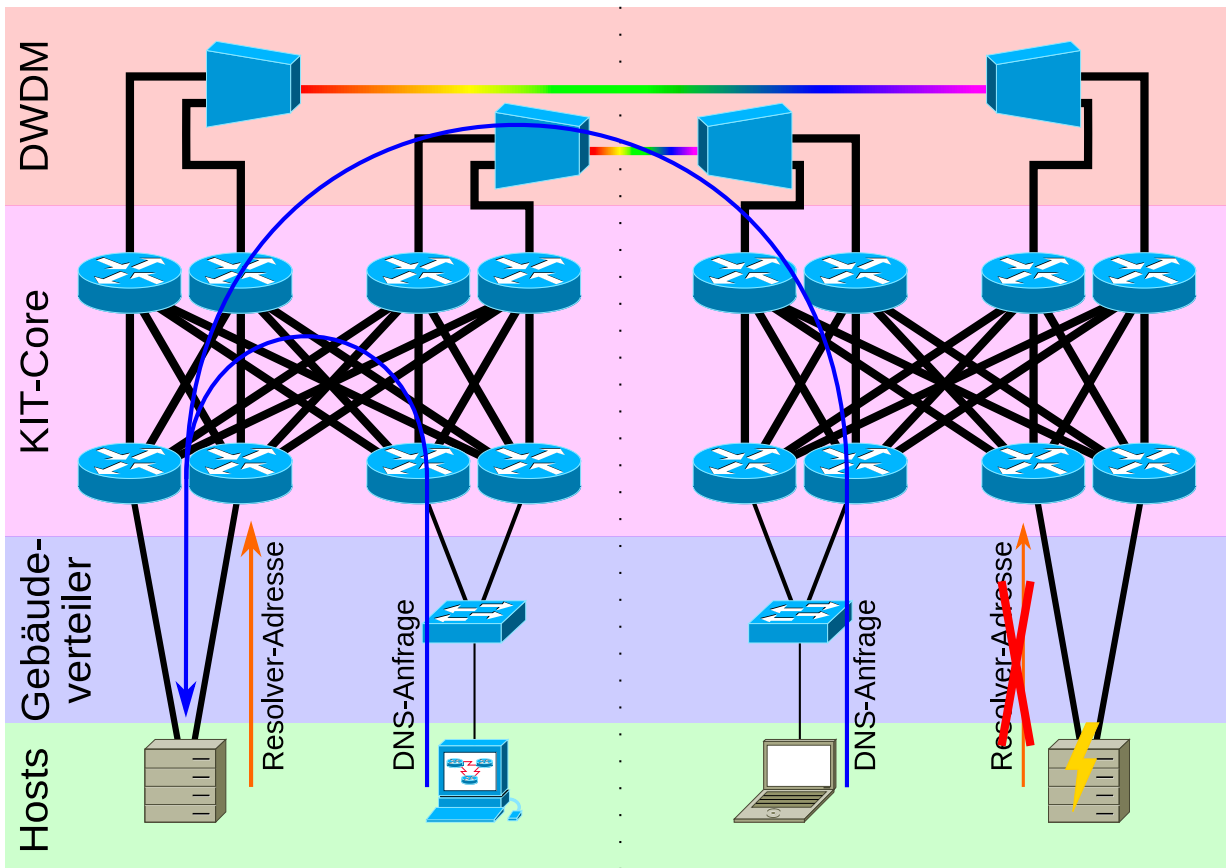


Abbildung 2:  
Im Fehlerfall

die Server die Resolver-Adressen über ein dynamisches Routing-Protokoll an die Router des KIT-Core und geben hierüber bekannt, dass sie auf Anfragen an alle Resolver-Adressen antworten. So ergeben sich im Netz für jede Resolver-Adresse verschiedene Wege, welche an verschiedenen Servern terminieren. Dies bringt zwei Vorteile: erstens wird ein System, das eine Resolver-Adresse anspricht, immer den kürzesten Weg nehmen, d.h. Anfragen gelangen immer zum aus Netzwerksicht nächstgelegenen Server. Zweitens wird im Wartungs- oder Fehlerfall der entsprechende Server die Resolver-Adressen nicht mehr gegenüber den Routern kommunizieren, so dass der Router die Anfragen nicht mehr an den entsprechenden Server weiterleitet. Alle Anfragen werden in diesem Fall zu den verbleibenden Servern geleitet. Diese Umschaltung passiert in kürzester Zeit, so dass kein Ausfall erkennbar ist. Diese Technik ist als Anycast bekannt (s. Abbildung 1 und 2).

### Resolver-Adressen am KIT

Mit Anycast muss man nun nicht mehr überlegen, welche Resolver-Adressen eingetragen werden müssen, da unabhängig von der verwendeten Adresse immer der nächstgelegene Server erreicht wird. Aus diesem Grund reichen zukünftig zwei Resolver-Adressen aus, weshalb wir uns entschieden haben, die jeweils zweiten Adressen der beiden Campus abzukündigen. Ab dem 1. Oktober 2022 wird es nur noch folgende DNS-Resolver-Adressen am KIT geben:

IPv6: 2a00:1398::1 und 2a00:1398::2  
IPv4: 129.13.64.5 und 141.52.3.3

Für einen Host mit IPv6 ist es ausreichend, die IPv6-Resolver-Adressen einzutragen. Die IPv4-Adressen sind nur bei IPv4-only-Hosts nötig. Die komplette Übersicht über die DNS-Server-Adressen ist auf den SCC-Webseiten zu finden:

[www.scc.kit.edu/dienste/dns-server](http://www.scc.kit.edu/dienste/dns-server)

### DNS resolvers

High availability of the DNS resolvers, which are used by almost all systems connected to the network, is required for a smooth operation. Client-side switching to the second "backup address" does not always work reliably, which is why a disruption of a resolver address can have significant impact on network operations. With the introduction of Anycast for these addresses, high availability is now ensured. Due to the deployment of Anycast, two of the four current DNS resolver addresses will be discontinued in October 2022. Please use 2a00:1398::1 and 2a00:1398::2, or 141.52.3.3 and 129.13.64.5 respectively for IPv4-only systems.

## Service Management in der European Open Science Cloud

Die Entwicklung und Etablierung eines zuverlässigen und effizienten Managements von IT-Dienstleistungen sind für den Erfolg der modernen Forschung von großer Bedeutung. In der European Open Science Cloud (EOSC) werden über 1,7 Millionen Wissenschaftlerinnen und Wissenschaftlern in Europa Daten und Dienste zur Verfügung gestellt, um ihre Forschung mit dem Schwerpunkt auf offener Wissenschaft und offenen Daten zu unterstützen und zu verbessern. Die Bereitstellung von Diensten, insbesondere auf europäischer Ebene, ist ein komplexes Unterfangen, das von dem von der EU geförderten Projekt EOSC-hub vorangetrieben wird. Ein wesentlicher Teil des Projekts besteht darin, die passenden Hierarchien und Richtlinien für ein prozessorientiertes Management der angebotenen IT-Dienstleistungen zu schaffen. Dieser Artikel befasst sich mit den Prozessen Incident und Service Request Management sowie dem Change Management, für die das SCC die Entwicklung leitet.

*Isabella Bierenbaum, Pavel Weber, Jos van Wezel*

### Einleitung

Die Entwicklung und der Aufbau einer verlässlichen und effizienten Verwaltung von IT-Services ist eine kontinuierliche Anstrengung für jedes Rechenzentrum und macht auch vor dem SCC nicht halt. Die formale Beschreibung und der Einsatz von Prozessen wurde ursprünglich von Unternehmen mit großen IT-Abteilungen praktiziert und in wissenschaftlichen Rechenzentren von Universitäten und Forschungsinstituten zunächst nur wenig eingesetzt, obwohl einige, darunter das SCC, von Anfang an die wichtigsten Teile des IT-Servicemanagement-Workflows erfolgreich implementieren konnten.

Verschiedene Veränderungen in der wissenschaftlichen IT-Landschaft haben dazu geführt, dass auch hier die Bedeutung einer überprüfbar und damit den Erwartungen entsprechenden Nutzung

von IT-Diensten wichtiger geworden ist. Dies beinhaltet unter anderem gestiegene Erwartungen an die Servicequalität und vermehrte Abhängigkeiten von der IT-Infrastruktur, eine größere Bedeutung der IT für den Erfolg moderner Forschung und Lehre sowie die Erweiterung und der Ausbau der IT-Infrastrukturen in alle Richtungen und damit einhergehende erhöhte Kosten der IT-Einrichtungen. In der Vision der European Open Science Cloud<sup>2</sup> (EOSC), in der Daten und Dienste allen Wissenschaftlern in Europa universell zugänglich gemacht werden, kommen alle diese Aspekte zusammen.

Das SCC setzt seine Erfahrungen in der Bereitstellung und dem Betrieb zuverlässiger Services mit Hilfe strukturierter Verfahren im EOSC-hub Projekt ein (s. SCC-News 1/2019, S. 20) und entwickelt gemeinsam mit Partnern die Workflows hinter den Serviceangeboten im

EOSC-Portal<sup>3</sup>. Über das EOSC-Portal werden Dienste präsentiert und können dort von jedem Forscher und jeder Forscherin in Europa bezogen werden. Das Portal ist der Hauptzugang zum EOSC-Servicekatalog und kumuliert die Dienste einzelner Labors und Forschungseinrichtungen bis hin zu großen paneuropäischen Infrastrukturen wie EGI und EUDAT-CDI.

Die kooperierenden Institutionen und Projekte in der EOSC werden durch eine Service-Governance- und Managementstruktur unterstützt, die auf dem FitSM<sup>4</sup> Prozessmanagement-Modell basiert. Der FitSM Standard kommt bekannten Frameworks wie ISO / IEC 20000 und ITIL, dem best practice Framework für das IT-Servicemanagement, sehr nahe und wurde vom Projekt EOSC-hub für die verteilte Verwaltung der Dienste in der EOSC ausgewählt.

Das Projekt EOSC-hub baut die Basis-Infrastruktur der EOSC auf, die eine Betriebsarchitektur auf zwei Ebenen vorsieht. Die Services des EOSC core bieten Funktionen an, auf deren Basis alle Benutzerservices aufbauen können und die eine EOSC-Föderation ermöglichen. Diese Services finden sich im **EOSC-hub Portfolio**. Die Benutzer-orientierten Dienste werden hier **EOSC Service Portfolio** genannt, während sie seit der Gründung einer Governance-Struktur der EOSC

The development and establishment of reliable and efficient management of IT services are of great importance for the success of modern research. In the European Open Science Cloud (EOSC), data and services are made available to over 1.7m scientists in Europe in order to support and improve their research with the focus on open science and open data. The delivery of services, especially on a European scale, is a complex undertaking that is led by the EU-funded project EOSC-hub<sup>1</sup> that builds the precursor of the IT infrastructure of the forthcoming EOSC. A fundamental part of the project consists of setting up the proper hierarchies and guidelines for the process-oriented management of the IT services. This article talks about the incident and service request management and the change management for which SCC leads the development.

<sup>1</sup> [www.eosc-hub.eu/](http://www.eosc-hub.eu/), EOSC-hub receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 777536

auch als **EOSC Exchange** (also EOSC-Börse) bekannt sind.

Der folgende Artikel beschreibt zwei der wichtigsten operativen Servicemanagement-Prozesse, an deren Aufbau und Durchführung das SCC im Rahmen des Projekts EOSC-hub wesentlich beteiligt ist: das Change Management und das Incident und Service Request Management.

## Das EOSC-hub Change Management

### Ziel und Aufgabe des EOSC-hub Change Managements

Ziel und Hauptaufgabe des EOSC-hub Change Managements (im Folgenden kurz als CHM bezeichnet) ist es, produktive Services vor negativen Auswirkungen bei Veränderungen (engl.: Changes) zu schützen. Dies beinhaltet etwa den EOSC Marktplatz<sup>5</sup>, einen zentralen Zugangspunkt zu den durch die EOSC angebotenen Services.

Ein Beispiel für solch eine „Änderung“ ist das Aufspielen eines Software-Updates an einem der beteiligten IT-Services. Aufgabe des CHM ist es, in diesem Fall zu verhindern, dass das Update absehbare negative Folgen für den betroffenen Service selbst, aber auch für davon abhängige Services und Nutzer hat. Wo dies nicht möglich ist, versucht das CHM im Rahmen seiner Möglichkeiten das Risiko für schädliche Auswirkungen zu minimieren.

„Changes“ umfassen jedoch mehr.

Sie können im Prinzip alle Elemente betreffen, die in der Konfigurationsdatenbank aufgeführt sind und beinhalten nicht nur das Aufspielen einer neuen Software, sondern etwa auch den Austausch von Hardware. Das CHM ist daher eng verbunden mit dem Konfigurationsma-

nagement, einem anderen Prozess im Service-Management-System (SMS) des Projekts EOSC-hub.

### Zuständigkeit des CHM und der Beitrag des SCC

In dieser Anfangsphase der EOSC erstreckt sich die Zuständigkeit des Change Managements hauptsächlich auf die Services im EOSC-hub Portfolio (auch einfach „Hub Portfolio“ genannt). Änderungen an diesen Diensten müssen formal im CHM angekündigt und beantragt werden. Das CHM dient ebenfalls dazu, neue Funktionalitäten dieser Dienste an den IT-Betrieb und Servicebetreiber in der EOSC zu kommunizieren und zu dokumentieren, um die Koordinierung von Integrationsaufgaben und die weitere Entwicklung von Services EOSC-weit zu erleichtern.

Das SCC hat einen großen Anteil am Aufbau des Change Managements innerhalb des EOSC-hub Projektes, an der Entwicklung und Definition der zugrundeliegenden Richtlinien und Prozeduren sowie deren Implementierung in ein Servicemanagement-Tool (JIRA<sup>6</sup>). Zudem ist das SCC an der Planung und Implementierung sowie am Betrieb des CHM von Beginn an beteiligt.

### Implementierung des CHM

Da das gesamte SMS des EOSC-hub Projektes dem FitSM-Modell folgt, liegt

dieses auch dem CHM zugrunde. Die darin enthaltenen Anforderungen, Rollen und Prozeduren wurden aber wo nötig verändert und angepasst, um der konkreten Situation im Hub zu entsprechen. Die Arbeitsabläufe innerhalb des CHM sind in einem auf Basis von JIRA implementierten Servicemanagement-Tool realisiert. Registrierte Nutzer, in der Regel die Service Owner der entsprechenden Services, können dort einen Request for Change (RfC) in Form eines sog. JIRA-Issues eröffnen und so den CHM-Prozess starten. Wichtige Angaben beim Öffnen eines RfC sind dabei die Dringlichkeit, die Reichweite sowie das Risiko eines geplanten Changes. Daher muss der Service Owner vor jedem Request eine Risikobewertung (engl. Risk Assessment) durchführen und dann, neben anderen Informationen, auch einen Risk Level für die Änderung angeben.

Jeder RfC wird von einem Change Owner durch alle Schritte seines Lebenszyklus hindurch betreut. Dieser Change Owner beurteilt auch, ob das Change Advisory Board (CAB), eine Gruppe aus Beratern und Entscheidern, die in die Prozesse auf Betreiberebene involviert sind, eingeschaltet werden muss, um die beantragten Changes vor ihrer Umsetzung eingehender auf mögliche Auswirkungen und Risiken zu überprüfen. Dies geschieht gegebenenfalls in Rücksprache mit allen beteiligten Parteien und falls notwendig externen Experten. Des Weiteren

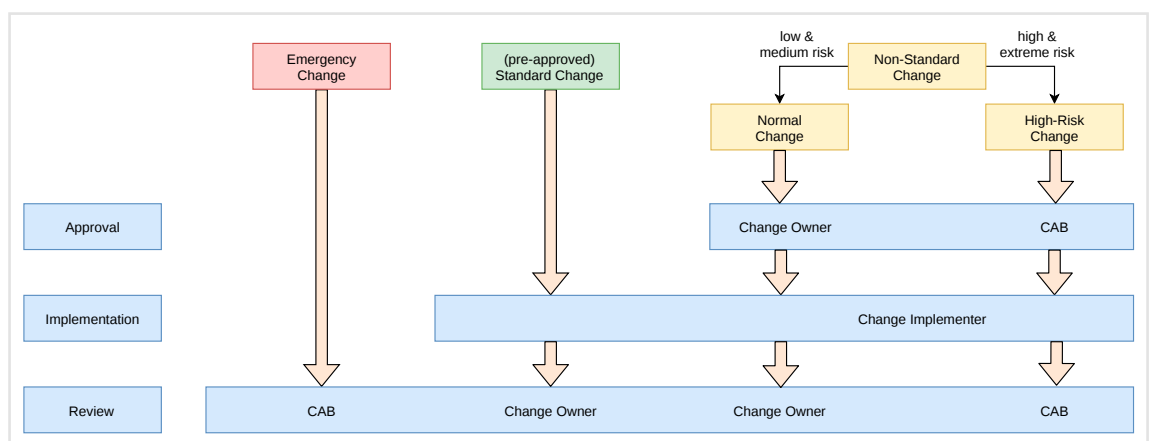


Abbildung 1: Arten von Changes im CHM. Vereinfacht dargestellt sind die wesentlichen Schritte des Verfahrens sowie die beteiligten verantwortlichen Personen in deren jeweiligen Rollen.

<sup>2</sup> Realising the European Open Science Cloud: [ec.europa.eu/research/openscience/pdf/realising\\_the\\_european\\_open\\_science\\_cloud\\_2016.pdf](https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf)

<sup>3</sup> [www.eosc-portal.eu](https://www.eosc-portal.eu)

<sup>4</sup> [www.fitsm.eu](https://www.fitsm.eu)

<sup>5</sup> [marketplace.eosc-portal.eu/](https://marketplace.eosc-portal.eu/)

<sup>6</sup> JIRA ist ein Issue-Tracker-System: [www.atlassian.com/software/jira](https://www.atlassian.com/software/jira)



begutachtet das CHM die Umsetzung der Changes und archiviert die RfCs.

### Mögliche Arten von Changes

Das CHM kennt drei Arten von Changes, die verschiedenen Situationen Rechnung tragen. Abbildung 1 skizziert die wesentlichen Schritte und verantwortlichen Personen.

- Emergency Changes:** Tritt ein Notfall ein, wie etwa ein Sicherheitsvorfall oder ein Hardwareausfall mit größeren Auswirkungen, muss der Service Owner unter Umständen sofort reagieren und Änderungen vornehmen, ohne auf Rückmeldungen von außen zu warten. Für solche sogenannten Emergency Changes kann daher erst nach der Implementierung aller notwendigen Maßnahmen ein RfC im CHM eröffnet werden, um die Änderung zu dokumentieren.
- Standard Changes:** Wiederkehrende Changes der gleichen Art, deren Einschätzungsprofil immer gleich ist, z.B. Software-Updates, können als Standard Changes registriert und als solche in einer Liste zur wiederholten Verwendung eingetragen werden. Dies erfordert eine einmalige Begutachtung und Genehmigung durch das CAB. Diese Changes sind damit für die Zukunft „im Voraus genehmigt“.
- Non-Standard Changes:** Sie bilden den Regelfall. Je nach Risk Level werden sie weiter unterschieden in Normal Changes oder High-Risk Changes. Während Normal Changes alleine vom Change Owner betreut werden können, müssen High-Risk Changes vom CAB begutachtet und genehmigt werden.

### CHM und die föderierte Umgebung des EOSC-hub

Eine Besonderheit für das Change Management des EOSC-hub liegt darin, dass es Services in einer föderierten Umgebung verwaltet und mit den dort auftretenden Besonderheiten umgehen muss. Nicht nur das Projekt EOSC-hub, auch die einzelnen Projektmitglieder besitzen eige-

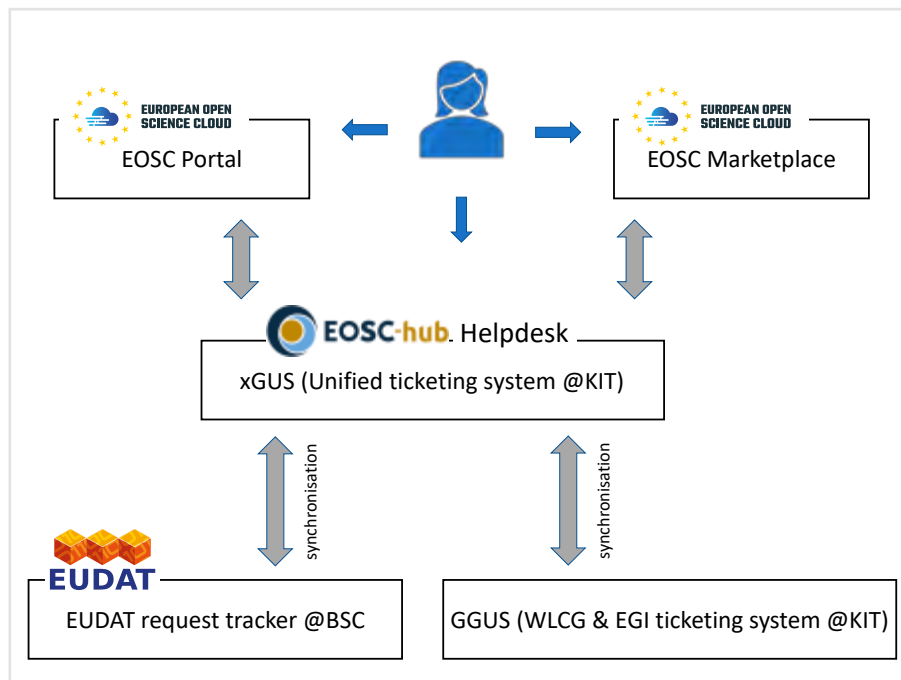


Abbildung 2: EOSC-hub Helpdesk System

ne lokale Change Management Prozesse, für die sie verantwortlich sind. Es gibt Dienste, die von einem der Projektmitglieder betrieben, aber vom gesamten EOSC-hub vertrieben und genutzt werden. Dies erfordert es, die lokalen Prozesse in das zentrale Service-Management-System zu integrieren, da bei einer Änderung an diesen „föderierten“ Diensten beide Change Managements involviert sind. Zu den Aufgaben von EOSC-hub gehört es daher auch, Regelungen und Übereinkommen mit den Projektpartnern zu finden, wie dieser wechselseitige Prozess ohne Schaden für die beteiligten SMS und deren Nutzer definiert werden kann.

### Zukunftsansicht und Pläne

Ein Ziel in der Fortentwicklung der EOSC und des damit verbundenen EOSC Portals und Marktplatzes ist es, zunehmend externe Serviceanbieter zu motivieren, dort ihre Services anzubieten. Diese Anbieter werden unterschiedlich ausgereifte SMS und damit auch Change Managements mit sich bringen. Falls die Services neue Dienste umfassen, die in die Zuständigkeit des EOSC-hub Portfolios fallen, müssen auch diese in das CHM aufgenommen, das heißt in das bestehende System

mit seinen Prozeduren und Richtlinien integriert werden. Wegen der starken Abhängigkeit des Change Managements vom Konfigurationsmanagement wird dies auch die lokalen Konfigurationsdatenbanken betreffen.

### Incident und Service Request Management im EOSC-hub

Um die Verwaltung der von EOSC-Nutzern eingereichten Serviceanfragen, die Bearbeitung von Störungen und die darauffolgende Wiederherstellung der Infrastruktur für geregelte Serviceabläufe zu ermöglichen, wurde im EOSC-hub der Incident und Service Request Management Prozess (ISRM) etabliert. Er definiert eine Reihe von Richtlinien für das Management von Incidents (Störungen) und Service Requests (Serviceanfragen), die beispielsweise eine durchschnittliche Reaktionszeit, die Struktur und Rollen der Support-Einheiten und die Bedingungen für eine Eskalation festlegen. Er definiert auch eine Reihe von Verfahren, die detaillierte Schritt-für-Schritt-Anweisungen für verschiedene Arten von Abläufen enthalten, wie zum Beispiel:

- Wie man mit Vorfällen und Serviceanfragen umgeht
- Wie und wann man einen Vorfall (Ticket) eskaliert
- Wie man eine neue Support-Einheit erstellt
- Wie man die Rolle für die verschiedenen Support-Ebenen des Helpdesks zuweist
- Wie man einen Vorfall oder eine Verschlechterung des Dienstes meldet

Der ISRM-Prozess liefert auch klare Anforderungen an die Helpdesk-Software und trägt zur Definition der Entwicklungs-Roadmap sowie zur Definition der Known Error-Datenbank und der Schnittstellen zu anderen Betriebsprozessen wie Change- oder Konfigurationsmanagement bei.

### Implementierung des Helpdesks im EOSC-hub

Der Helpdesk-Prozess des EOSC-hubs wurde im Laufe des Projekts in einem einheitlichen Ticketing-System implementiert, um die Anfragen zu verschiedenen

Service- oder Ressource-Anbietern aus unterschiedlichen e-Infrastrukturen zu verwalten.

Anfänglich lag der Schwerpunkt auf der Integration der Helpdesks von zwei großen e-Infrastrukturen, die zum EOSC-hub beitragen, nämlich EUDAT und EGI. Das SCC hat zu dieser Aktivität beigetragen, indem es die seit Jahren etablierte und immer weiterentwickelte xGUS-Helpdesk-Technologie (s. SCC-News 2/2014, S. 21) als Front-End für den Helpdesk des EOSC-hub zur Verfügung stellte und die vollständige Synchronisierung zwischen dem Helpdesk des EOSC-hub und denen von EUDAT und EGI ermöglichte. Vollständige Synchronisierung bedeutet, dass in xGUS ankommende Tickets automatisch über API-Schnittstellen an die Helpdesks von EUDAT und EGI weitergeleitet werden und weitere Änderungen am Ticket innerhalb dieser Systeme im EOSC-hub Helpdesk angezeigt werden.

Darüber hinaus wurde der Helpdesk des EOSC-hub mit dem EOSC-Portal und dem EOSC-Marktplatz verbunden und in die Authentifizierungsinfrastruktur (AAI) der EOSC integriert, so dass der Benutzer mehrere Möglichkeiten hat, eine Anfrage entweder über Portale oder direkt an den Helpdesk zu stellen (s. Abbildung 2).

### Die Struktur der Support-Einheiten

Um der Herausforderung einer komplexen verteilten Architektur des EOSC mit vielen Service Providern gerecht zu werden, benötigt der zukünftige EOSC Helpdesk eine mehrstufige Struktur der Support-Einheiten. Die vom SCC vorgeschlagene Struktur wurde erfolgreich umgesetzt und wird ständig um neue Support-Einheiten erweitert. Eine Übersicht über diese Helpdesk-Ebenen und die wichtigsten Support-Einheiten ist in Abbildung 3 dargestellt.

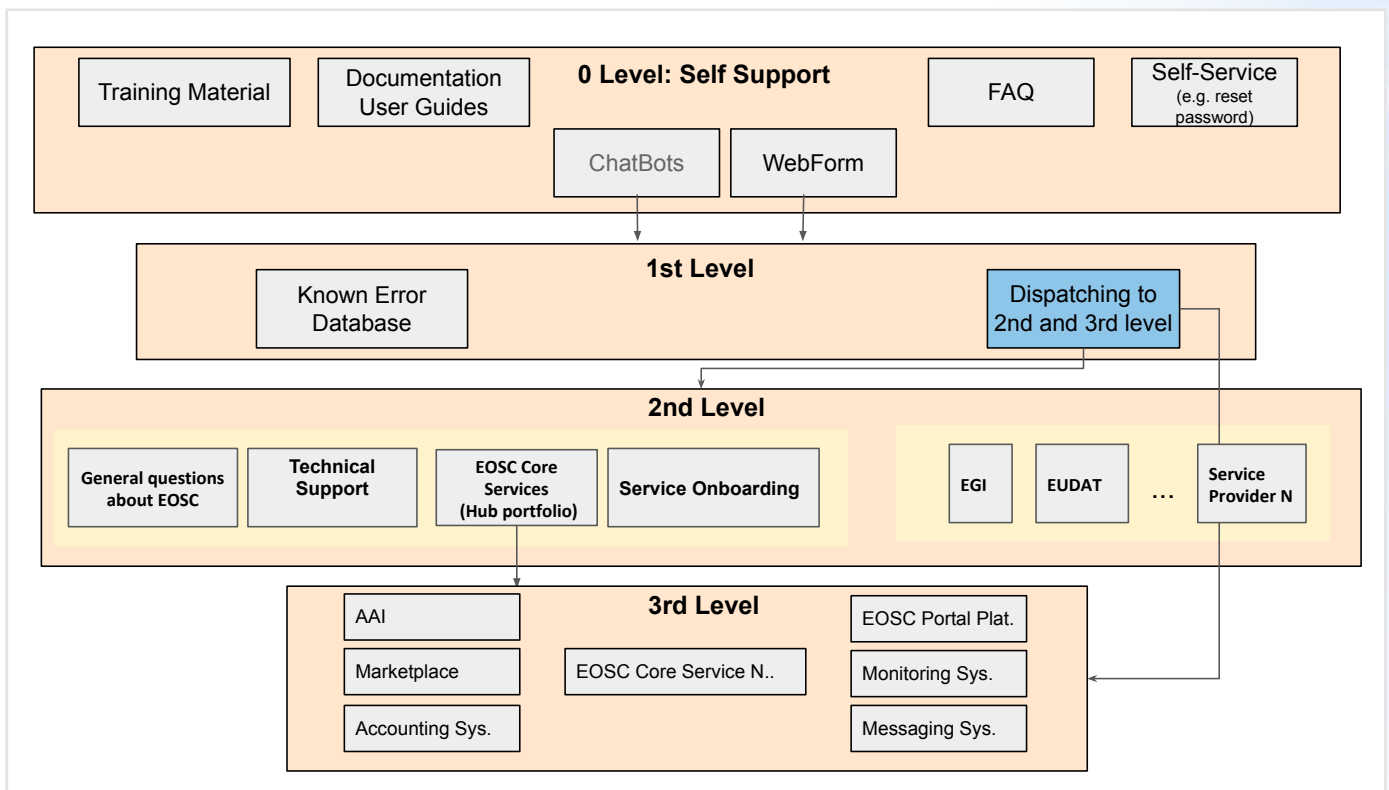


Abbildung 3: Überblick über die Helpdesk-Ebenen und die wichtigsten Support-Einheiten im EOSC-hub

Derzeit sind vier Hauptebenen (Level) des Helpdesk des EOSC-hub definiert:

- **0. Level:** Self-Support auf Basis der bereitgestellten Dokumentation, FAQ usw.
- **1. Level:** Lösung grundlegender Probleme mit Hilfe der Known Error Database oder Allgemeinwissen. Weitergabe von Tickets an entsprechende Einheiten entweder an den 2. oder die 3. Level.
- **2. Level:** Gezielte technische und allgemeine Unterstützung durch die EOSC-Dienstleister oder durch zentrale Support-Einheiten, die von erfahrenen EOSC-hub-Experten geleitet werden.

- **3. Level:** Umfassende Unterstützung auf Code-Ebene durch Entwickler oder Service-Eigentümer. Diese Ebene ist ausschließlich für EOSC-hub vorgesehen.

#### Zukunftsaussicht und Pläne

Das etablierte Helpdesk-System im Projekt EOSC-hub sowie der ISRM-Prozess werden in den nächsten EOSC-Projekten weiter ausgebaut. Die künftige Arbeit wird sich auf die Integration von Helpdesk-Systemen verschiedener Infrastrukturen, die zur EOSC beitragen, und auf das Hinzufügen weiterer Support-Einheiten für die EOSC-Dienste konzentrieren. Außerdem

ist die Integration des ISRM-Prozesses mit dem EOSC-Konfigurationsmanagementsystem geplant, um den Change Management Prozess zu verbessern, das SLA-Management durchzuführen und automatisierte Abläufe für die Weiterleitung und Eskalation der Tickets zu implementieren.

## bwSync&Share – der Landesdienst migriert zu Nextcloud

Zum April 2020 hat das Betriebsteam die komplette Installation des Landesdienstes bwSync&Share auf die neue Software-Basis Nextcloud umgestellt. Mit dieser Migration ergeben sich neue Möglichkeiten der Nutzung und ein besseres Bedienerlebnis für die Anwenderinnen und Anwender. Der sichere und zuverlässige Betrieb des Landesdienstes wird so weiterhin gewährleistet. Zudem ergeben sich neue und erweiterte Nutzungsszenarien – bis hin zur zukünftigen Verwendung von schon bestehenden oder selbst entwickelten Apps.

*Matthias Knoll, Klaus Scheibenberger, Claudius Laumanns*

### bwSync&share – Nextcloud als passende Fortentwicklung

Nextcloud, als eine Lösung am Markt, die sich einer breiten Anwendung erfreut und seit einiger Zeit auch in föderierten Umgebungen gut einsetzbar ist, erweist sich als solide Basis für die weitere Entwicklung des Dienstes bwSync&Share. Zudem kennen einige Nutzerinnen und Nutzer die Software schon von anderen Einrichtungen, was vielen den Ein- oder Umstieg erleichtert. Die bekannte Nutzung des Dienstes sowohl über das Web-Interface als auch mit Desktop-Clients (Linux, Windows) und Mobile Apps (iOS, Android) bleibt erhalten. Insbesondere

der Desktop-Client zeigt sich in einer schlankeren, effektiveren Form und läuft stabil im Hintergrund. Die moderne Weboberfläche erleichtert die komfortable Bedienung im Webbrowser. Freigaben auf Dokumente oder Bilder sind schnell erstellt. Dateien und Ordner können Anwender entweder direkt an andere am Dienst Teilnehmende und an Gäste freigeben. Durch die weiterhin vorhandene Integration von OnlyOffice können Dokumente und Dateien wie gewohnt direkt im Browser, auch mit anderen zusammen, bearbeitet werden.

Der Dienst ist weiterhin unter [bwsync-candshare.kit.edu](https://bwsync-candshare.kit.edu) erreichbar. Auch alle

bwSync&Share-Clients verbinden sich über diesen Endpunkt mit den Synchronisations-Servern.

### Erweiterte Funktionalität

Gegenüber der vorherigen Softwarelösung der Fa. Powerfolder bietet Nextcloud eine deutlich erweiterte Funktionalität. Neue Ordner und Ordnerstrukturen können im Datei-Explorer innerhalb des Synchronisationsordners angelegt bzw. geändert und weitere Ordner der lokalen Dateistruktur über den Client zur Synchronisation hinzugefügt werden. Auch mit den mobilen Applikationen ist die Handhabung von Ordnern und Datei-

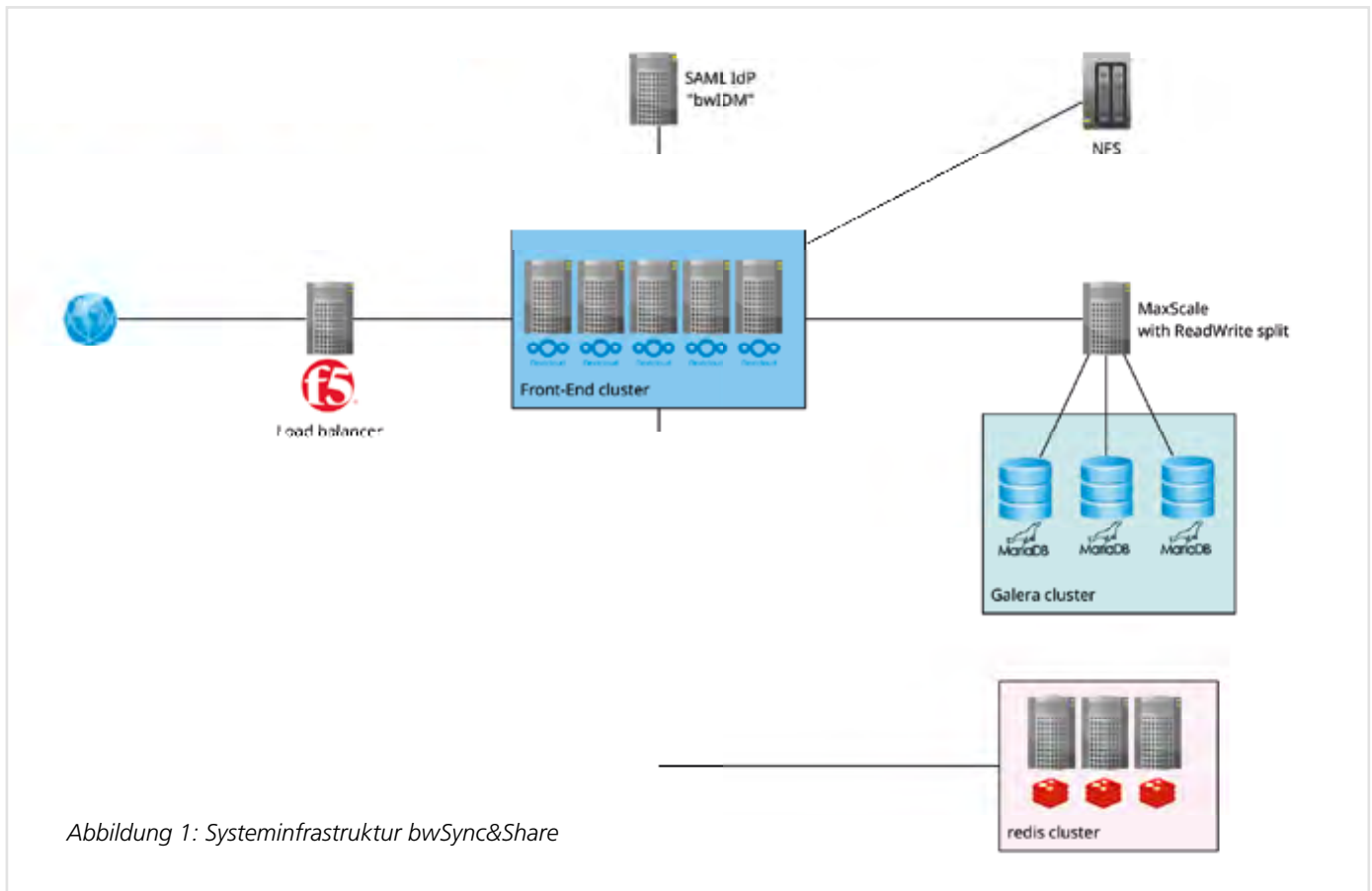


Abbildung 1: Systeminfrastruktur bwSync&Share

en nun wesentlich flexibler. So können diese Objekte beispielsweise einfach umbenannt oder in der Ordnerstruktur verschoben werden.

Wie bisher ist es möglich, sowohl Dokumente als auch Ordner mit Nutzenden von teilnehmenden Einrichtungen als auch mit Gästen zu teilen. Dies ist nun nicht mehr nur auf Ordner der obersten Ebene beschränkt, sondern auch für Unterordner sowie einzelne Dateien möglich. Die Einladungen zu freigegebenen Ordnern müssen nicht mehr explizit angenommen werden (so die Voreinstellung). Die Ordner, die auf die lokale Festplatte synchronisiert werden sollen, können im Client gezielt ausgewählt werden. Damit bleibt der lokale Speicherplatzverbrauch unter Kontrolle. Der Besitz von Ordnern und Dokumenten ist in den Profileinstellungen auf eine andere Person mit Speicherkontingent übertragbar, um z.B. die kontinuierliche Arbeit von Arbeitsgruppen auch bei einem Wechsel von Mitgliedern sicherzustellen.

Eine weitere Option, Ordner und Dokumente öffentlich zu teilen, ist der Versand eines Freigabe-Links. Dieser Link berechtigt zum Download von Dateien. Wenn der Link mit der Berechtigung „Hochladen und Bearbeiten erlauben“ erstellt wurde, können auch Dateien im Ordner abgelegt werden. Passwortschutz und Ablaufdatum für den Link erhöhen die Sicherheit. Soll der Zugriff auf den Ordner nur ausgewählten Personen erlaubt werden, muss dieser Ordner explizit für sie geteilt werden. Neu ist die Berechtigung „Dateien ablegen“. Sie erlaubt Dateien abzulegen bzw. nur hochzuladen, ohne dass die Inhalte des Ordners sichtbar sind.

Falls eine Ende-zu-Ende Verschlüsselung aus Vertraulichkeitsgründen erforderlich ist, kann diese mit einer zusätzlichen Cloud-Verschlüsselungslösung (z.B. Cryptomotor) über den Desktop-Client eingerichtet werden (s. FAQ unter [help.bwsyncandshare.kit.edu](http://help.bwsyncandshare.kit.edu)). Eine in bwSync&Share, d.h. in die Software Nextcloud, integrierte Ende-zu-Ende-

Verschlüsselung ist aber bereits im Bestadium der Nextcloud-Entwicklung vorhanden und wird im Dienst verfügbar gemacht, sobald eine stabile Version verfügbar ist.

Eine weitere interessante Funktionalität von Nextcloud sind die sog. Groupfolder. Das sind Ordner, die nicht an das Konto einer Person gebunden sind. Damit lassen sich Dateizugriffsrechte, wie sie beispielsweise für Projektordner benötigt werden, besser abbilden. Hierfür sind im bwSync&Share-Kontext noch Anpassungen der Software notwendig.

### Technische Strukturen

Die derzeit rund 45.000 Nutzerinnen und Nutzer greifen über insgesamt fünf Nextcloud-Knoten auf den Dienst zu (Abbildung 1). Auf diesen Knoten wird das Web-Frontend direkt in PHP-Code ausgeführt. Die Kombination aus Apache2 mit Multi-Processing- (MPM) und PHP-Modulen beschleunigt zudem



die Auslieferung an die Desktop-Clients. Über den vorgeschalteten Load Balancer F5 Big-IP kann so zu Spitzenzeiten jeder dieser Frontend-Knoten um die 2.000 Verbindungen aufrecht halten, da Nextcloud bei Nutzerinteraktionen auf der Webseite parallele Verbindungen mit dem jeweiligen Apache-Knoten aufbaut. Ebenso bauen die Clients bei Synchronisierungsvorgängen parallele Verbindungen auf, um die Geschwindigkeit zu optimieren. Das macht zwar die Oberfläche und Clients performant, stellt aber einige Anforderungen an die eingesetzten Apache-Module und deren Konfiguration. Um die bis zu 10.000 Verbindungen sicher zu handhaben, belegt beispielsweise Apache-MPM prefork RAM- und Speicher-Ressourcen für eine höhere Anzahl von gerade aktiven Verbindungen vor. Die vorbereiteten Prozesse warten auf Verbindungen und werden beim Zustandekommen einer Verbindung in einem eigenen Prozess isoliert. So beeinflusst ein einzelner fehlgeschlagener Prozess andere Verbindungen nicht. Außerdem lassen sich mit den passenden Einstellungen des MPM prefork (z.B. Anzahl der voreinstellbaren Tochterprozesse oder maximale Anzahl paralleler Apache-Prozesse) die Ressourcen stabil bereitstellen und deren Auslastung planen.

Ein Datenbank-Galera-Cluster aus vier MariaDB-Servern mit vorgeschaltetem Maxscale-Load Balancer bearbeitet die Datenbank-Anfragen der Frontend-Server. Im Unterschied zur Vorgängerlösung ist in der Datenbank die komplette Verzeichnisstruktur aller Nutzenden und deren Metadaten abgebildet, so dass die Software für die Anzeige der Ordner- und Dateiliste im Webbrowser zunächst nur die Datenbank konsultiert, also noch keine Operationen auf dem Dateisystem notwendig sind. Des Weiteren speichert die Installation vorkompilierte PHP-Dateien und temporäre Daten für jeden Nutzenden in einem Redis-Cluster, um das Framework weiter zu beschleunigen (Abbildung 1).

## Migration zu Nextcloud

### Dienstprozesse

Neben den rein intern-funktionalen Aspekten der Software wie beispielsweise die Nutzeroberfläche und deren Performance, die Effizienz der Handhabung von Daten oder die Zusammenarbeit mit anderen Teilnehmern, gibt es weitere Punkte, die den Integrationsgrad des Dienstes für den Nutzer und die nutzenden Einrichtungen wesentlich mitbestimmen. Das sind u.a. die Einbettung in eine Identitätsföderation (hier bwIDM) für das Zugriffs- und Identitätsmanagement sowie die technische Umsetzung von Dienstprozessen zur Steuerung der Nutzung. Hierzu

#### EINIGE ECKDATEN ZUR MIGRATION:

- Inhaltsdaten: Ca. 60 Terabyte Daten, ca. 65 Mio. Dateien und Ordner
- Metadaten: Ca. 40.000 Konten, 57.000 Einladungen
- Die abschließende Zusammenführung/Kopplung von Nutzerdaten und Metadaten wurde mit 72 parallelen „file scan threads“ auf 5 Nextcloud Servern durchgeführt.

gehört z.B. der Einladungsprozess, der sicherstellt, dass die eingeladene Person die Nutzungsbedingungen bestätigt und sich somit am Dienst registriert. Diese Aspekte wurden intensiv mit der Fa. Nextcloud erörtert, die die erforderlichen Anpassungen und Erweiterungen der Software sehr professionell durchgeführt hat.

### Technische Migration

Die tatsächliche Umstellung der Plattform wurde so geplant, dass die gespeicherten Inhaltsdaten (Dateien) und die personenbezogenen Daten (Account-Daten und Einladungen an Teilnehmer) zu einem bestimmten Datum automatisiert, d.h. ohne Eingriff durch die Nutzer, in die neue Nextcloud-basierte Umgebung übertragen wurden. Nach der Umstellung konnten sich alle Nutzenden wieder über [bwsyncandshare.kit.edu](https://bwsyncandshare.kit.edu) verbinden und ihre bisherigen Daten und Einladungen vorfinden.

Dazu musste eine Abbildung der Datenbankstrukturen der Vorgänger-Software auf die Nextcloud-Strukturen (u.a. Metadaten) erarbeitet werden. Auf SCC-Seite mussten Mechanismen geschaffen werden, um die ca. 60 Terabyte gespeicherten Daten auf die von der Nextcloud-Software geforderten Datenstrukturen im Filesystem abzubilden. Die dafür entwickelten Migrationsautomatismen wurden gemeinsam von der Fa. Nextcloud und Mitarbeitern des SCC in mehrfachen Testläufen seit November 2019 getestet und iterativ verbessert. Bis zum Jahresende konnten nicht alle Tests rechtzeitig abgeschlossen werden, da auch einige Anpassungen in der Software, z.B. für die Provisionierung von Nutzern, erst mit dem neuen Release ab Mitte Januar 2020 umgesetzt werden konnten. Daher wurde die zum Ende des Jahres 2019 geplante Umstellung um drei Monate, auf Ende März 2020 verschoben.

Die so vorbereitete, automatisierte Umstellung erfolgte schließlich von Freitag 27.03.2020 bis Montag 30.03.2020. Die Migration dauerte insgesamt 52 Stunden. Nach anschließenden Überprüfungen und Tests konnte der Betrieb wie geplant am 31.03.2020 wieder aufgenommen werden.

#### bwSync&Share – the federal state service migrates to Nextcloud

As of April 2020, the operating team has converted the complete installation of the federal state service bwSync&Share to the new software basis Nextcloud. This migration opens up new usage options and a better user experience. The secure and reliable operation of the service will thus continue to be guaranteed. In addition, new and expanded usage scenarios will be available - including the future use of already existing or self-developed apps.

# Digitalisierte Rechnungsverarbeitung am KIT

Um die vorwiegend manuellen und teilweise sehr arbeits- und zeitaufwändigen Prozesse in der Rechnungsverarbeitung zu verbessern und zu vereinheitlichen, führte das KIT eine digitalisierte Rechnungsverarbeitung ein. 2018 startete das Finanzmanagement des KIT das Projekt „Einführung Automatisierte Rechnungsbearbeitung“ und evaluierte und implementierte gemeinsam mit dem SCC die dazu nötige Softwarelösung. Im März 2020 wurde die Lösung für Beschäftigte des KIT produktiv gesetzt.

*Gerald Helck, Mark Herzog*

## Rechnungsbearbeitung am KIT wird digital

Das KIT als große Bildungs- und Forschungseinrichtung verarbeitet jährlich ca. 135.000 Eingangsrechnungen. Die Rechnungen gehen zu ca. 60% als E-Mail mit PDF-Anhang ein, der Rest in Papierform. Im bisherigen Prozess wurden alle eingehenden Rechnungen in der Finanzbuchhaltung von Hand erfasst. Bevor die Rechnungen bezahlt werden können, müssen die Freigaben der jeweiligen Verantwortlichen aus den Organisationseinheiten (OE) vorliegen. Dabei wurden am Großforschungsbereich (GFB) und Universitätsbereich (UB) unterschiedliche Verfahren genutzt. Am GFB wurden nach der Buchung der Rechnung alle zur Zahlungsfreigabe erforderlichen Genehmigungen über Papierformulare eingeholt, die von den Zeichnungsberechtigten der beschaffenden OE unterschrieben werden mussten. Zur Umsetzung der Rechnungsfreigaben im SAP-System waren KIT-spezifische Programmiererweiterungen notwendig. Im UB erfolgte die Rechnungsfreigabe über ein dezentral erstelltes und unterschriebenes Kontierungsblatt. Das von den Zeichnungsberechtigten unterschriebene Formular wurde zusammen mit der Eingangsrechnung an die Finanzbuchhaltung weitergeleitet. Verbesserungspotential gab es vor allem hinsichtlich der vielen Medienbrüche im Gesamtprozess, der teilweise langwierigen Genehmigungsvervollständigung und des Zugriffs auf den Originalbeleg am UB.

Um diese Prozesse zu verbessern und zu vereinheitlichen, lag es nahe, die Einführung einer automatisierten Rechnungsbe-

arbeitung (AREB) zu prüfen. 2018 führte das Finanzmanagement des KIT (FIMA) das Projekt „Einführung Automatisierte Rechnungsbearbeitung“, geleitet von Jürgen Stehmer, durch. Das SCC unterstützte bei technischen Fragen und übernahm die Integration in die bestehende ERP-Systemlandschaft. Im März 2020 wurde die Lösung für Beschäftigte des KIT produktiv geschaltet.

## Prozessschritte der automatisierten Rechnungsverarbeitung

Viele Firmen bieten seit geraumer Zeit Software zur digitalen Rechnungsverarbeitung an. Vereinfacht dargestellt unterscheiden alle Lösungen die folgenden Prozessschritte:

### Digitalisierung der Rechnung

Alle Papierrechnungen werden gescannt und die ausgelesenen Daten mit SAP-Stamm- und -Bewegungsdaten plausibilisiert, um eine hohe Erkennungsqualität zu gewährleisten.

### Archivierung

Zusätzlich werden das Originaldokument sowie wichtige andere Zusatzinformationen in einem Dokumentenarchiv gespeichert. Das digitalisierte Dokument ist später auf Knopfdruck am Bildschirm verfügbar und eine langwierige Suche in Aktenordnern oder im E-Mail-Postfach entfällt. Der Zugriff auf den optisch abgelegten Rechnungsbeleg war bisher nur im Großforschungsbereich möglich.

### Rechnungsdaten auslesen

Über eine OCR-Erkennung<sup>1</sup> werden nach der Digitalisierung die Rechnungspositions-

und Kopfdaten wie Absender, Empfänger, Rechnungsnummer und Beträge ausgelesen und in einem Rechnungseingangsbuch gespeichert. Eine manuelle Erfassung der Rechnungsdaten durch die Rechnungsprüfer ist nicht mehr erforderlich. Das Rechnungseingangsbuch ist der zentrale Bestandteil der AREB, der Informationen bereits vor der eigentlichen Verbuchung speichert und den aktuellen Freigabestatus der Rechnungen dokumentiert.

### Validierung

Die Informationen, die durch die OCR-Erkennung nicht bzw. nicht korrekt erkannt wurden, werden manuell durch die Sachbearbeiterinnen und Sachbearbeiter in der Buchhaltung korrigiert und vervollständigt. Die Validierung stellt auch fest, ob Stammdaten der Lieferanten, z.B. neue Bankverbindungen, aktualisiert werden müssen.

### Genehmigung der Eingangsrechnung

In diesem Prozessschritt prüfen Verantwortliche der beschaffenden OE die Rechnung bezüglich Preis und Menge sowie sachlicher Richtigkeit. Zur Beschleunigung des Informationsaustauschs werden idealerweise elektronische Workflowverfahren genutzt.

### Buchung

Die Finanzbuchhaltung gibt die Rechnung frei, sobald alle Genehmigungen vorliegen.

### Vorteile und Anwendung der Lösung

Diese, in einer Softwarelösung integrierten, digitalen Prozesse haben große Vorteile. Mit der OCR-Erkennung und der Workflowtechnik reduzieren sich z.B. die Erfassungs- und Weiterbearbeitungszeiten

<sup>1</sup> Die Schrift aus dem Digitalisat wird in reinen Text umgewandelt, so dass Programme damit umgehen können.

enorm. Zahlungsziele und Skontofristen lassen sich leichter einhalten. Der gesamte Rechnungsprüfungsprozess ist über das Rechnungseingangsbuch nachverfolgbar und transparent. Durch die Kombination von automatisierter Datenübernahme und Validierung werden Eingabefehler reduziert und somit die Datenqualität erhöht. Die Anwenderinnen und Anwender können über einen Zugang zum Dokumentenmanagementsystem digitalisierte Rechnungsbelege anzeigen. Papierbelege müssen nicht mehr aufbewahrt werden – die freiverdende Lagerfläche kann für andere Zwecke genutzt werden.

Die implementierte Anwendung ist als SAP-AddOn in die SAP-ERP-Lösung des KIT eingebettet, so dass Genehmigende, die über einen SAP-Nutzerstamm verfügen, über das SAP-Portal auf den Web-Arbeitsplatz mit allen von ihnen zu bearbeitenden Rechnungen zugreifen. Für Anwenderinnen und Anwender ohne SAP-Nutzerstamm steht ein eigenständiges Web-Portal zur Verfügung. Die Nutzenden werden automatisch über den anstehenden Arbeitsvorrat informiert. Neben vielen Funktionen, die im direkten Zusammenhang mit der Freigabe von Rechnungen stehen, kann am Web-Arbeitsplatz eine Vertretung bei Abwesenheit hinterlegt werden. Darüber hinaus können die Anwenderinnen und Anwender über das elektronische Rechnungseingangsbuch nach Rechnungen der eigenen OE bzw. selbst bearbeiteten Rechnungen suchen.

### Digitalized Invoice Processing at KIT

In order to improve and standardize the mainly manual and partly very labor-intensive and time-consuming processes in invoice processing, KIT introduced a digitized invoice processing system. In 2018, KIT's financial management started the project "Introduction of Automated Invoice Processing" and evaluated and implemented the necessary software solution together with SCC. In March 2020, the solution went live for KIT employees.

### Wissenswertes

Am Ende eines umfangreichen Auswahlprozesses fiel die Wahl auf das Produkt xSuite der Fa. xSuite Group. Nach der Erstellung des Feinkonzepts und der Installation der Software bis Juli 2019 wurden in der zweiten Jahreshälfte intensive Tests aller relevanten Geschäftsvorfälle durch den Fachbereich durchgeführt. Parallel erweiterte xSuite die Standardlösung um KIT-spezifische Anforderungen. Im Laufe des Januars und Februars 2020 bereiteten Mitglieder des Projektteams zunächst FIMA und dann die Organisationseinheiten an mehreren Schulungsterminen auf die neue Abwicklung vor. Mitte März konnte die Lösung als SAP-AddOn in die SAP-ERP-Lösung eingebettet und produktiv gesetzt werden.

Bereits im Vorfeld hat das SCC die bestehende IT-Infrastruktur um einen XFlowCapture-Server ergänzt, der gescannte Belege in vom Finanzbuchhaltungssystem verarbeitbare Belege umsetzt. Um auch die Anzeige der digitalisierten Belege im Dokumentenarchiv komfortabel zu gestalten, wurde mit dem docPresenter ein neues Web-basiertes Zugriffsverfahren implementiert. Mit der Einführung des docPresenters ist auch eine Umstellung des Speicherformats erforderlich, so dass die Ablage nicht mehr im tif- sondern im pdf-Format erfolgt. Zur Optimierung der Anzeige werden auch alle archivierten Belege des GFB aus den letzten Jahren nach pdf umgesetzt.

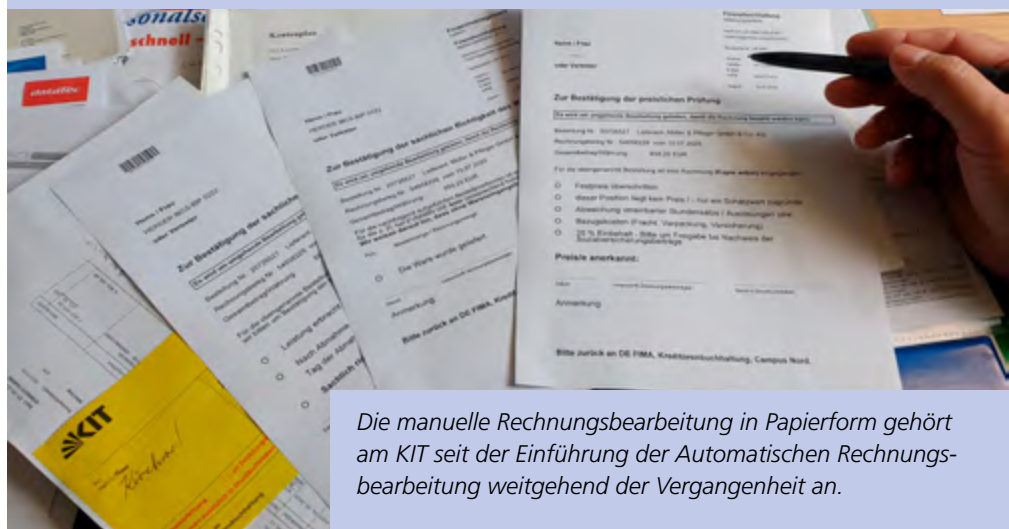
Das SCC hat die Einführung der Lösung technisch unterstützt. Unter anderem mit Anpassungen der SAP-Berechtigungen, dem Transport technischer Objekte im Entwicklungs-, Qualitäts- und Produktivsystem, dem Einbau von SAP-Korrekturen und Einspielen neuer Produktreleases von xSuite. Im Zuge der Synchronisation der Geschäftsprozesse wurden Anpassungen in den SAP-Programmen vorgenommen, z.B. Unterdrückung des Druck bisheriger Freigabeformulare am GFB, Verzicht auf KIT-spezifische Genehmigungsarten am GFB, Nutzung von Bestellungen ohne erforderlichen Wareneingang am UB.

### Fazit und Ausblick

Die Produktivsetzung von xSuite war gerade abgeschlossen, als sich die Corona-Pandemie verschärfte und die meisten Mitarbeitenden des KIT ins Homeoffice gingen. Gerade noch rechtzeitig konnten die elektronischen Workflows mit digitaler Genehmigung aktiviert und so der Rechnungsprüfungsprozess aufrechterhalten werden. Die Übermittlung von handschriftlich abgezeichneten Formularen per Hauspost wäre unter den gegebenen Umständen nicht oder nur mit immenser Verzögerung möglich gewesen.

Es war eine große Kraftanstrengung vieler am Prozess beteiligter Personen, diese weitreichenden Prozessänderungen umzusetzen. Dank ihnen verlief die Einführung erfolgreich, und das KIT konnte seinen Zahlungsverpflichtungen nahtlos weiter nachkommen. Der Rechnungseingangsprozess am KIT ist bedingt durch seine besondere Organisationsform als Forschungsuniversität in der Helmholtz-Gemeinschaft sehr komplex. Im Produktionsbetrieb konnten weitere Probleme identifiziert und Verbesserungsmöglichkeiten aufgezeigt werden.

Ein wichtiges Ziel, das mit dem Abschluss des Projekts ebenfalls erreicht wurde, ist die Harmonisierung und Vereinheitlichung des Rechnungsprüfungs-, Archivierungs- und Genehmigungsprozesses am gesamten KIT. Ein einheitlicher Prozess reduziert die Komplexität, er lässt sich leichter an veränderte Gegebenheiten anpassen und ist deutlich weniger fehleranfällig. An der Optimierung der Prozesse wird permanent gearbeitet.



Die manuelle Rechnungsbearbeitung in Papierform gehört am KIT seit der Einführung der Automatischen Rechnungsbearbeitung weitgehend der Vergangenheit an.

## Vorbereitung der Ginkgo-Softwarebibliothek für AMD-Grafikprozessoren

Das Hochleistungsrechnen auf Grafikprozessoren wurde in den letzten Jahren vom Chiphersteller NVIDIA dominiert, obwohl auch andere Hersteller sehr leistungsfähige Hardware bieten. Die auf Commodity-Geräte, aber auch auf die HPC-Welt abgestimmte Entwicklungsumgebung von NVIDIA ist dabei ein Hauptgrund, diese Chips einzusetzen. Doch es bewegt sich etwas auf dem Markt, denn bald werden Exascale-Systeme mit AMD-Chips ausgestattet sein. Die zugehörige Entwicklungssoftware nimmt durch ihren Namen schon vorweg, dass diese Systeme bald angesagt - HIP - sein könnten. In diesem Innovationsschritt spielt die am SCC entwickelte Ginkgo C++-Softwarebibliothek für High Performance Computing ihre Stärken aus: Sie konnte dank der flexiblen Architektur vom Ginkgo-Entwicklerteam mit sehr gutem Ergebnis um ein HIP-Backend erweitert werden.

*Hartwig Anzt, Tobias Ribizel*



In den letzten Jahren war GPU-Computing, das Hochleistungsrechnen auf Grafikprozessoren (GPU), eine One-Man-Show des Chipherstellers NVIDIA: Alle relevanten GPU-beschleunigten Supercomputer basierten auf NVIDIA GPUs. Dies erstreckt sich vom Piz Daint Supercomputer in der Schweiz über den Titan in Oak Ridge, USA, bis hin zu seinem Nachfolger Summit, dem bis vor Kurzem weltweit schnellsten Supercomputer.

NVIDIAs Dominanz auf dem GPU-Markt für Hochleistungsrechner wird unter anderem von der benutzerfreundliche CUDA<sup>1</sup>-Entwicklungsumgebung getragen, die GPU-Computing neben Supercomputern auch auf Desktop-PCs, Laptops und eingebetteten Systemen zugänglich macht.

CUDA ist unbestritten die wichtigste GPU-Programmiersprache und auch das ForHLR II-Nachfolgesystem HoreKa am KIT (siehe Seite 35) wird mit NVIDIA-GPUs ausgerüstet sein. Doch trotz NVIDIAs

starker Marktstellung gibt es Anzeichen, dass diese Vorherrschaft zu einem Ende kommen könnte: In einer Ankündigung des US-amerikanischen Exascale Computing Projekts wurde bekannt, dass zwei der drei geplanten neuen Supercomputer mit AMD-GPUs ausgestattet werden sollen, wohingegen keine neuen Systeme mit NVIDIA-GPUs geplant sind. Es ist zu erwarten, dass dieser Trend von anderen Systemherstellern aufgegriffen und eine zunehmende Verbreitung von AMD-GPUs zu sehen sein wird.

Ein bedeutender Unterschied zu AMDs vorherigen Anstrengungen im HPC-Markt ist, dass ihre Hardware-Fortschritte nun mit einem Software-Ökosystem unterstützt werden, das vergleichbar mit NVIDIAs CUDA-Entwicklungsumgebung ist. In der Tat weist die darin enthaltene Programmiersprache „HIP“<sup>2</sup> nicht nur große Ähnlichkeiten mit der CUDA-Programmiersprache auf, HIP-Code kann beinahe nahtlos auf NVIDIA-Architekturen kompiliert und ausgeführt werden. Dabei konvertiert ein Source-to-Source Skript den HIP-Code in CUDA-Code, der im Anschluss mit Hilfe von NVIDIAs nvcc Compiler kompiliert wird. Mit zunehmendem Interesse an AMD GPU-Clustern will auch das Ginkgo-Team seine Software für die HIP-Programmierung vorbereiten.

Generell ist die Portierung eines Softwarepakets auf eine neue Hardware-Architektur ein mühsames Unterfangen. Wie in den SCC-News 1/2019 beschrieben, lag ein starker Fokus auf der Erweiterbarkeit der Ginkgo-Softwarebibliothek. Dieses innovative Design ermöglichte es jetzt, Ginkgo mit moderatem Aufwand um ein HIP-Backend zu erweitern. Den Grundstein dafür legt die konsequente Trennung zwischen der High-Level-Struktur und Low-Level-Kernels für sämtliche in Ginkgo implementierten Algorithmen (siehe Abb. 1).

Die Algorithmen bestehen zu großen Teilen aus dem Aufruf hardware-spezifischer Kernels, was wiederum bedeutet, dass neue Backends hinzugefügt und entfernt werden können, ohne dass Änderungen an der High-Level-Implementierung der Algorithmen nötig sind oder die Ausführung auf anderen Backends beeinflusst wird. AMD unterstützt die Übersetzung von CUDA- in HIP-Code mit dem hipify-Skript, das bei der Portierung von Ginkgo zu HIP unserer Erfahrung nach einen

<sup>1</sup> **CUDA** (Compute Unified Device Architecture) is a parallel computing platform and application programming interface (API) model created by NVIDIA (Quelle wikipedia)

<sup>2</sup> AMD: HIP programming guide [rocmdocs.amd.com/en/latest/Programming\\_Guides/HIP-GUIDE.html](http://rocmdocs.amd.com/en/latest/Programming_Guides/HIP-GUIDE.html)



Großteil des Codes erfolgreich übersetzen konnte. Nichtsdestotrotz unterstützt CUDA einige Features, die in HIP nicht oder unvollständig umgesetzt sind, etwa cooperative groups, sodass für einen Teil von Ginkgos CUDA-Code eine manuelle Übersetzung und Implementierung dieser Features in HIP nötig war (siehe [1]).

Während dieser Implementierung des HIP-Backends für Ginkgo konnten wir starke Überlappungen zwischen unserem HIP- und CUDA-Code feststellen – teilweise unterschieden sich die Dateien nur in hardware-spezifischen Parametern wie der Größe der Warps/Wavefronts oder Caches. Da der primäre Unterschied zwischen CUDA und HIP somit im Aufruf der Kernel-Funktionen liegt, konnten wir sämtliche identischen Implementierungen in eine „common“ Codebasis auslagern, deren Kernel-Templates während der Kompilierung für das spezifische GPU-Backend instanziiert werden.

Mit dieser Strategie konnten wir eine Code-Vervielfältigung vermeiden und so insgesamt den Code-Umfang verringern. Zusammen wurden mit der Integration des HIP-Backends etwa ein Drittel des Codes in „common“, sowie je ein Drittel in CUDA und HIP-spezifischen Code aufgeteilt (Abb. 2). Mit dem fertiggestellten HIP-Backend für Ginkgo will das Development-Team<sup>3</sup> nicht zuletzt die Nutzung von HIP für NVIDIA-Architekturen betrachten. Insbesondere stellt sich dabei die Frage, welchen Performance-Effekt die Nutzung des HIP-Backends anstelle von nativem CUDA-Code in Ginkgo als High-performance-Bibliothek für Lineare Algebra hat.

Wie zu erwarten war, ist der native CUDA-Code im Durchschnitt etwas schneller, allerdings sind die Performance-Unterschiede zu HIP kleiner als 5%. Dies zeigt, dass AMDs HIP-Ökosystem ein GPU-Programmiermodell liefert, das erfolgreich auf diversen Grafikkartenarchitekturen eingesetzt werden kann.

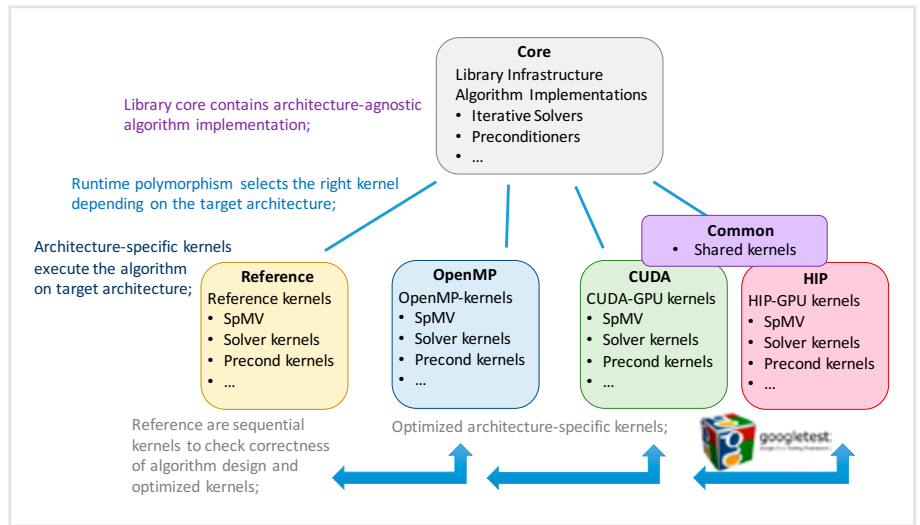


Abbildung 1: Architektur der Ginkgo Softwarebibliothek

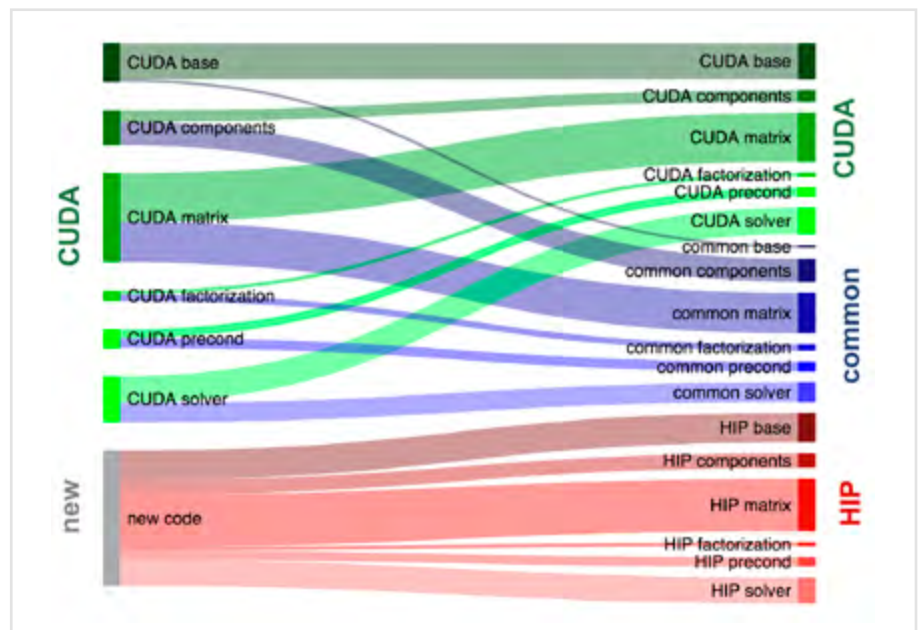


Abbildung 2: Restrukturierung der Ginkgo Softwarebibliothek zur Integration eines HIP-backends

**Weitere Informationen:**

[1] Tsai, Cojean, Ribizel, Anzt: Preparing Ginkgo for AMD GPUs -- A Testimonial on Porting CUDA Code to HIP submitted to HeteroPar 2020, preprint available:

[arxiv.org/abs/2006.14290](https://arxiv.org/abs/2006.14290)

[2] Blog-Artikel: [bssw.io/blog\\_posts/porting-the-ginkgo-package-to-amd-s-hip-ecosystem](https://bssw.io/blog_posts/porting-the-ginkgo-package-to-amd-s-hip-ecosystem)

<sup>3</sup> H. Anzt, T. Cojean, F. Göbel, T. A. Grützmaier, P. Nayak, T. Ribizel, Y.-H. Tsai

# Porting Ginkgo Software Library to AMD's HIP Ecosystem



Over the years, general purpose computing on graphics processors (GPUs) was a one-man-show run by NVIDIA: At least since NVIDIA's GPUs started serving as the powerhouse behind the Piz Daint supercomputer in Switzerland and the Titan Supercomputer at the Oak Ridge National Laboratory in the US, virtually all GPU-accelerated supercomputers relied on NVIDIA technology, including Titan's successor, the Summit supercomputer, until recently ranked No. 1 in the peak performance metric. Featuring a user-friendly programming interface, NVIDIA's CUDA ecosystem has grown to a de-facto standard in GPU programming on supercomputers as well as desktop stations and embedded devices.

However, even though the CUDA ecosystem is still dominating the market, and KIT just announced the ForHLR II-successor HoreKa will feature the newest NVIDIA Ampere GPU architecture, there is more movement in the market than ever before: AMD is gaining enough traction to challenge the NVIDIA hegemony with their ROCm ecosystem. The strategy behind the ROCm effort is different from AMD's previous attempts, as it puts significant focus on the software ecosystem and offers with the HIP language a programming interface very similar to the NVIDIA CUDA ecosystem. However, the HIP language takes an even more aggressive step by providing the possibility to compile HIP code for both AMD and NVIDIA architectures. This is realized by utilizing source code conversion in

combination with NVIDIA's nvcc compiler. Furthermore, AMD is making the complete ROCm ecosystem and HIP compiler open source. But not only the software landscape is moving: The US Exascale Computing project already announced that two of the three planned Exascale supercomputers will feature AMD GPUs, while none of them will feature NVIDIA components. The fastest supercomputers in the world relying on AMD GPUs will likely not remain without consequences, and we may expect an increasing number of Supercomputers featuring AMD GPUs.

Anticipating the rise of AMD GPU clusters, we are interested in preparing our software to back end the HIP language. Generally, porting a software package to a new hardware architecture using a different programming model is a cumbersome challenge. However, the sophisticated design of the Ginkgo library, which keeps extensibility as a primary design principle, allows us to create a Ginkgo back end in the HIP ecosystem at moderate effort. This is possible as the Ginkgo library architecture consistently separates the algorithm implementations from the architecture-specific back ends, see Figure 1. The algorithms themselves are unable to execute without the hardware-specific kernels, and adding a back end does not require any changes to the core algorithms. This allows the addition and removal of hardware back ends as required without impacting the library's functionality on other back ends.

To simplify the porting to the HIP ecosystem, AMD provides a "hipify" script that converts CUDA code to HIP code. When creating a HIP back end for Ginkgo, our experience is that most of the code can be handled by this script, dramatically

reducing the porting effort. As the HIP ecosystems currently lacks some technical functionality available in the CUDA ecosystem, there remains a small portion of the code that we have to port manually, e.g. cooperative groups, see [1].

During the creation of the HIP back end for Ginkgo, we have noticed a significant level of similarity between the CUDA and the HIP implementation of the kernels. Often, the kernels are identical except for hardware-specific parameters like warp size or cache configuration. With the kernel launch expression often being the only significant difference, we created a "common" code base containing the template kernels that are configured during the compilation process for the specific back ends. This strategy avoids code duplication and reduces the size of the code. In the realization of the HIP back end, we rearranged Ginkgo's code base to split the code into 1/3 "common" code, 1/3 HIP-specific code, and 1/3 CUDA-specific code, see Figure 2.

After having completed a production-ready HIP back end for Ginkgo, we are interested in the possibility to compile HIP kernels for NVIDIA architectures. Specifically, with Ginkgo being designed as a high performance linear algebra library, we are interested in the performance penalty we pay when compiling HIP code for an NVIDIA back end instead of writing native CUDA code.

As expected, the native CUDA code typically runs slightly faster. However, we acknowledge that the performance differences are within a 5% range, which demonstrates that AMD succeeds in providing a cross-platform GPU programming interface.

## Die Helmholtz Metadata Collaboration Plattform

Mit der Helmholtz Metadata Collaboration Plattform startete Ende 2019 ein wichtiges Themenfeld des Helmholtz Inkubators "Information & Data Science", der die Expertise der Helmholtz-Zentren zusammenführt und das Thema „Information & Data Science“ über die Grenzen von Zentren und Forschungsbereichen hinweg gestaltet. Übergeordnetes Ziel der Plattform ist, die qualitative Anreicherung von Forschungsdaten durch Metadaten langfristig voranzutreiben, die Forschenden zu unterstützen – und dies in der Helmholtz-Gemeinschaft und darüber hinaus umzusetzen.

*Thomas Jejkal, Sabine Chelby, Andreas Pfeil, Rainer Stotzka*

### „FAIRe“ Forschungsdaten durch Metadaten

Heute gibt es wohl kaum noch Forschende, die gänzlich ohne digitale Forschungsdaten auskommen. Bei der Datenerfassung werden dabei oft die Grenzen des Machbaren erreicht, um hochwertige Informationen aus den zu untersuchenden Forschungsobjekten zu extrahieren. Dies betrifft alle sechs Forschungsbereiche der Helmholtz-Gemeinschaft. Ob es sich um Patientendaten (Gesundheit), um das Spektrogramm eines zu untersuchenden Materials (Materie), oder um Strahlungsmessungen aus dem Weltall (Luftfahrt, Raumfahrt und Verkehr) handelt, spielt dabei keine Rolle. Für die Forschenden sind diese Daten die Essenz ihrer wissenschaftlichen Erfolge. Fakt ist aber auch, dass Wissenschaftlerinnen und Wissenschaftler typischerweise 80% ihrer Zeit mit der Aufbereitung genau dieser Daten für ihre Analyse verbringen. Neben der eigentlichen Identifikation und Sammlung relevanter Daten ist oft auch die Transformation und Zusammenführung von Daten aus verschiedenen Quellen für die Weiterverarbeitung notwendig. Somit bleibt den Forschenden noch etwa 20% ihrer Zeit für die inhaltliche Datenauswertung und den daraus resultierenden Erkenntnisgewinn. Dies hat zur Folge, dass Aufgaben wie die Dokumentation der Daten und ihre Aufbereitung für die Wiederverwendbarkeit in den Hintergrund treten und dadurch ein Großteil der Forschungsdaten nur für einen kleinen Teil Forschender nutzbar ist. Eine Veröffentlichung dieser Daten ist oft mit großem, nachträglichem Aufwand verbunden und

genießt daher in vielen Bereichen noch nicht denselben Stellenwert wie etwa die wissenschaftliche Veröffentlichung der Ergebnisse in Form eines Papers.



*(c) Sungya Pundir, Wikimedia Commons, CC By-SA 4.0*

Ein weltweit anerkannter Ansatz für nachhaltiges Forschungsdatenmanagement wurde im Jahr 2016 durch das Paper 'FAIR Guiding Principles for scientific data management and stewardship' [1] aufgezeigt. Darin sind vier Grundeigenschaften von Forschungsdaten definiert (Abbildung 1), die benötigt werden, um diese langfristig auffindbar, nachvollziehbar und wiederverwendbar zu machen. Neben der eigentlichen Auffindbarkeit (Findable) durch eindeutige und unveränderliche Bezeichner werden die Zugreifbarkeit (Accessible), etwa durch technische Maßnahmen und Lizenzen sowie die Interoperabilität (Interoperable) und Wiederverwendbarkeit (Reusable) durch den Einsatz von Metadaten beschrieben. Metadaten bilden Zusatzinformationen zu den eigentlichen Forschungsdaten und können je nach Nutzungsszenario unterschiedlich ausfallen. Für die Nachvollziehbarkeit sind etwa die Einstellun-

gen verwendeter Messgeräte oder die Parameter der für die Analyse verwendeten Algorithmen relevant, während für die Veröffentlichung von Forschungsdaten Informationen über beteiligte Forschungseinrichtungen und Zugriffsinformationen in Form von Lizenzen benötigt werden.

Ziel der Helmholtz Metadata Collaboration (HMC) Plattform ist es, Anleitung, Beratung und vor allem einfach benutzbare Werkzeuge zur Verfügung zu stellen, die Wissenschaftlerinnen und Wissenschaftlern den „FAIRen“ Umgang mit ihren Forschungsdaten ermöglichen und erleichtern.

### Die Struktur der HMC Plattform

Aus fachlicher Sicht ist die HMC in sechs verschiedene Metadata Hubs organisiert. Jeder Metadata Hub deckt dabei einen der sechs Forschungsbereiche Erde & Umwelt, Energie, Gesundheit, Materie, Luftfahrt, Raumfahrt und Verkehr und Information & Data Science ab. Die Vertreter eines Metadata Hubs agieren als Kontaktstellen für die Forschenden dieser Forschungsbereiche und die zentralen Einheiten der HMC. Diese sind durch das HMC Office repräsentiert, welches für die Projektkoordination und die projektübergreifende Vernetzung verantwortlich ist (Abbildung 2). Mit der Einrichtung eines Arbeitspaketes "FAIR Data Commons: Technologien und Prozesse" wird die Definition von allgemeingültigen Prozessen und Infrastrukturdiensten für die Helmholtz-weite, einheitliche Umsetzung der FAIR-Prinzipien vorangetrieben. Die im Arbeitspaket erstellten Basisdienste

werden, ergänzt durch einfach nutzbare Werkzeuge, allen Forschenden der Helmholtz-Gemeinschaft und darüber hinaus zur Verfügung gestellt. In den Metadata Hubs der Forschungsbereiche werden diese in fachspezifische Lösungen integriert. Dadurch können auf den organisatorischen und den technischen Ebenen einheitliche Repräsentationen der Forschungsdaten erreicht werden, die Kollaborationen innerhalb und zwischen den Forschungsbereichen wesentlich vereinfachen. Das KIT engagiert sich in der HMC mit der Koordination des Metadata Hubs Energie (Institut für Automation und angewandte Informatik, IAI) und der Mitarbeit im Arbeitspaket "FAIR Data Commons: Technologien und Prozesse" (SCC).



Abbildung 2: Organisationsstruktur der HMC mit Metadata Hubs für jeden Forschungsbereich und zentralen Einrichtungen.

**FAIR Data Commons**

Die Entwicklung im Rahmen von "FAIR Data Commons: Technologien und Prozesse" wird in enger Kooperation des Forschungszentrums Jülich und des SCC durchgeführt.

Das SCC bringt hierfür seine langjährige Erfahrung im Forschungsdaten- und Metadatenmanagement ein, wie beispielsweise:

1. Expertise aus dem Programm Supercomputing & Big Data und in den Data Life Cycle Labs
2. Erfahrungen in großen internationalen Forschungsdaten-Projekten, z.B. der European Open Science Cloud
3. Aufbau von Forschungsdateninfrastrukturen, z.B. LSDF und GridKa
4. Aufbau von speziellen Community-Repositoryn für verschiedene Forschungsbereiche
5. Entwicklung der Repository-Plattform KIT Data Manager [2]
6. Entwicklung neuer Technologien für Metadatenverwaltung und -nutzung, z.B. zur automatischen Metadatenextraktion
7. Daten- und Metadatenversionierung

Besonders im Umfeld der Research Data Alliance (RDA) sind Vertreter des SCC seit vielen Jahren maßgeblich bei der Erstellung des Konzepts von FAIR Digital Objects (siehe Textbox) beteiligt. Als Grundlage für die Umsetzung dieses Konzepts dient die Verwendung von Metadaten. Diese werden durch allgemeingültige Dienste und Prozesse mit den Forschungsdaten verknüpft und liefern so wertvolle Informationen für die Interpretation der Daten. Möchte man nun noch einen Schritt weiter gehen, etwa um eine automatische Durchsuchbarkeit von Forschungsdaten zu realisieren, benötigt man neben inhaltlichen Metadaten weitere, maschinenlesbare Metadaten. Dies sind zum Beispiel Informationen über Struktur und Inhalt von Daten und Metadaten sowie eindeutige und unveränderliche Bezeichner für eine

zeitlich uneingeschränkte Referenzierbarkeit der Forschungsdaten und assoziierter Metadaten. Durch die Realisierung einer automatischen Durchsuchbarkeit können Wissenschaftlerinnen und Wissenschaftler zum Beispiel automatisch Referenzdatensätze für ihre Forschung identifizieren und verwenden. Um hierfür den Grundstein zu legen, bietet die HMC durch die Etablierung von standardisierten und einheitlichen Methoden und Prozessen mit einem hohen Automatisierungsgrad. So entstehen notwendige Komponenten, um Forschungsdaten nachhaltiger und vielfältiger nachnutzen zu können und somit die Forschung in der Helmholtz-Gemeinschaft wesentlich zu unterstützen.

**FAIR Digital Objects**

Bei einem FAIR Digital Object handelt es sich um eine Darstellung eines digitalen Datensatzes, etwa einer Datei, welche die Anforderungen der FAIR-Prinzipien erfüllt. Hierzu muss es auffindbar (Findable), zugreifbar (Accessible), interoperabel (Interoperable) und wiederverwendbar (Reusable) sein. Metadaten und Standards spielen hier eine wichtige Rolle. Grundlage für die Umsetzung von FAIR Digital Objects ist die Verwendung von eindeutigen und unveränderlichen Bezeichnern auf allen Ebenen und eine maschinenlesbare Beschreibung sowohl von Daten als auch von Metadaten. Dies erlaubt unter anderem eine automatisierte Suche nach Forschungsdaten, die bestimmten Kriterien entsprechen, sowie deren Validierung. Den Forschenden ist es so einfach möglich, Vergleichsdaten zu finden und in ihre Arbeit einzubeziehen.

[1] [www.nature.com/articles/sdata201618](http://www.nature.com/articles/sdata201618)  
 [2] [github.com/kat-data-manager](https://github.com/kat-data-manager)

**Helmholtz Metadata Collaboration Platform**

The Helmholtz Metadata Collaboration (HMC) Platform tackles an important topic of the Helmholtz Incubator "Information & Data Science" that was launched in the end of 2019. Overarching goal of HMC is to foster the qualitative enrichment of research data with metadata. This goal will be achieved by developing and introducing novel technologies and services to support metadata management, training and consultancy services into the daily research. For the scientists, this will make the transition to FAIR research data even easier.



## Wieso manchmal weniger mehr ist – auch in der Wissenschaft!

In einer Zusammenarbeit zwischen dem SCC und dem Institut für Meteorologie und Klimaforschung (IMK) wurden neue Algorithmen zur Kompression von Simulationsdaten entwickelt. Die Algorithmen reduzieren den Speicherbedarf von strukturierten Daten ohne Genauigkeitsverluste mit einem durchschnittlichen Kompressionsfaktor von 2,6. Das hat Auswirkungen auf unser aller Leben: Wetter- und Klimavorhersagen können nämlich bei limitierten Speicherressourcen mit höherer Genauigkeit erstellt werden, um u.a. die Klimapolitik eines Landes zuverlässig zu unterstützen; der CO<sub>2</sub>-Ausstoß von Rechenzentren wird reduziert, weil mit der Speicherung von mehr Daten die Anzahl von Simulationen reduziert werden kann; die Bandbreitenausnutzung wird verbessert, da die auf Speichermedien zu schreibende Datenmenge reduziert wird; die Lebenszeit der Festplatten wird verlängert, indem ihre Ausnutzung verringert wird; und wissenschaftliche Simulationen können potentiell schneller berechnet werden, da mehr Daten im Hauptspeicher gehalten werden können.

Uğur Çayoğlu

### Wie Klimamodelle unser Leben beeinflussen

Viele von uns denken am Abend, wenn wir z.B. in der Tagesschau die Wettervorhersage sehen, nicht darüber nach, wie diese Vorhersagen erstellt werden. Sie ist so ein integraler Bestandteil unseres Lebens geworden, dass wir es uns nicht vorstellen können, einmal keine(!) Vorhersage für die kommenden Tage zu bekommen. Heutzutage erhalten wir sogar minutengenaue Vorhersagen auf unser Smartphone, wann der Regenschauer endlich aufhört, der uns gerade verzweifelt vor dem Schaufenster kleben lässt, weil es gefühlt die einzige trockene Stelle der ganzen Stadt ist. Die Geschwindigkeit, mit der sich die Wetter- und Klimamodelle weiterentwickeln, ist beeindruckend, insbesondere wenn man bedenkt, wie jung das Forschungsfeld ist und welche Infrastruktur nötig ist, um ein erfolgreiches Wetter- bzw. Klimamodell zu bilden.

Die ersten Wettermodelle basierten auf einfachen mathematischen Konzepten. Das Ziel war es, eine Vorhersage basierend auf der aktuellen Entwicklung des Wetters zu geben – und das so schnell wie möglich. Schließlich ist es nicht besonders hilfreich, wenn die "Vorhersage" erst nach dem vorherzusagenden Ereignis ankommt.

Deshalb haben Wissenschaftlerinnen und Wissenschaftler an ihren Standorten den aktuellen Zustand des Wetters aufgenommen und diese Daten mit einer neuen Entwicklung, dem Telegrafen, an andere Standorte weitergegeben. Das war die einzige Möglichkeit, beispielsweise dem Nachbarort Informationen weiterzugeben, bevor das Wetter bei diesen eintraf.

Der Telegraf ermöglichte zum ersten Mal die zentrale und schnelle Verarbeitung von Informationen und die Erfassung von größeren Mengen an Messwerten. Denn um das Wetter an einem Ort für einen bestimmten Zeitraum vorherzusagen, müssen Messwerte von mehreren Orten zu einem Zeitpunkt erfasst werden. Das Wetter in Deutschland ist abhängig von dem seiner Nachbarländer und umgekehrt, wodurch sich eine Kultur der Informationsweitergabe in der Wettervorhersage und den Klimawissenschaften zwischen den Ländern entwickelte.

Der erste Forscher, der sich ernsthafte Gedanken über die mathematische Modellierung des Wetters gemacht hat, war Vilhelm Bjerknes. Er analysierte am Anfang des 20. Jahrhunderts die Zusammenhänge der Fluid- und Thermodynamik und entwickelte die sogenannten primitiven Gleichungen (engl. „Primitive Equations“). Diese Gleichungen verwenden jedoch viele

Näherungen und werden heute nur noch für bestimmte Anwendungen benutzt. Aktuelle Modelle verwenden Modellgleichungen mit weniger Näherungen.

Später bieten Satelliten eine bahnbrechende Möglichkeit, Wetterbeobachtungen in die Vorhersagen einzuspeisen. Die erste nahezu globale Abdeckung wurde schließlich durch Wettersatelliten erreicht, deren Finanzierung John F. Kennedy im Rahmen seiner berühmten Rede am 25. Mai 1961 vor dem amerikanischen Kongress angekündigt hat<sup>1</sup>. Sie ging leider etwas unter, weil er im ersten Teil seiner Rede das Ziel verkündete, Amerika werde als erstes Land einen Menschen zum Mond und sicher wieder zurückbringen.

Bereits in diesen Anfängen zeichnete sich der Kampf der Klimaforschung gegen zwei Herausforderungen ab: Die schnelle Prozessierung der Daten und die nachhaltige Hinterlegung der Daten (siehe Beispiel in Abbildung 2) für zukünftige Vorhersagen bzw. die Verbesserung der Modelle.

### Klimamodelle stellen noch immer große Herausforderungen an moderne Hardware

Diese beiden Herausforderungen sind auch noch bis heute gültig. Klimaforschen-

de versuchen mit Hilfe von Künstlicher Intelligenz und besserer Ressourcennutzung die Auflösung der Modelle zu erhöhen, um die Klimapolitik bestmöglich zu unterstützen und Naturkatastrophen wie Hitzewellen und Dürreperioden besser vorherzusagen. Gleichzeitig generieren diese Simulationsläufe aber eine so große Datenmenge, dass das Institut für Meteorologie und Klimaforschung am KIT die mit Abstand größten Speicherkapazitäten des SCC benötigt.

Um diese Herausforderungen anzugehen, wurde am SCC gemeinsam mit dem IMK im Rahmen einer Doktorarbeit [1] das Potential der Klimadatenkompression erforscht. Das Ergebnis dieser Arbeit ist ein neues Kompressionsverfahren, welches im Schnitt einen 10% höheren Kompressionsfaktor für Klimadaten erreicht als der aktuelle Stand der Wissenschaft (siehe Abbildung 3). Der aktuelle Stand der Technik mit fpzip bietet einen höheren Durchsatz (Abbildung 1). Der Fokus dieser Arbeit liegt auf der Reduktion des Speicherplatzes wofür der Kompressionsfaktor entscheidend ist. Mit pzip wird im Schnitt ein Kompressionsfaktor von 2,6 erreicht, sodass mehr als die Hälfte des Speicherplatzes wieder zur Verfügung steht. Der Algorithmus wurde auch mit Daten anderer Wissenschaftsbereiche getestet und hat hier ebenfalls eine bis zu 30% bessere Kompression erreicht.

**Methodik**

Der Kompressionsalgorithmus ist ein sogenanntes vorhersagebasiertes Kompressionsverfahren. Hierbei werden die einzelnen Datenpunkte im Datensatz nach einem festgelegten Muster durchlaufen. Während dieser Traversierung wird für jeden einzelnen Datenpunkt eine Vorhersage berechnet. Diese Vorhersagen berücksichtigen die Position und Distanz der vorhergegangenen Datenpunkte zum aktuellen Datenpunkt. Sie kann durch a priori-Wissen über die Zusammenhänge einzelner Variablen unterstützt werden. Ein Beispiel für ein solches a priori-Wissen ist, dass die zeitliche Entwicklung der

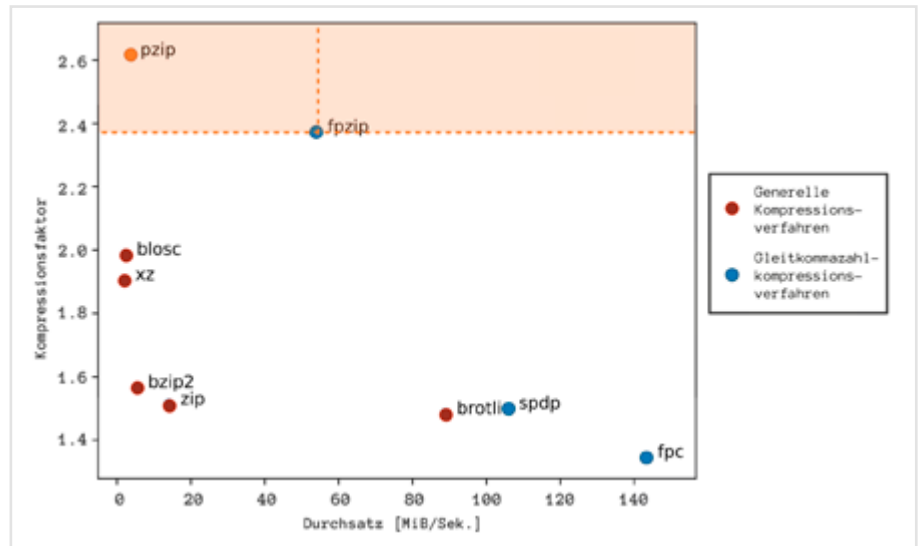


Abbildung 1: Vergleich von aktuellen Kompressionsverfahren mit dem am SCC und IMK entwickelten pzip Kompressionsalgorithmus bzgl. Kompressionsfaktor und Durchsatz bei der Kompression [1].

Temperatur am Äquator „stabiler“ ist als am Nord- bzw. Südpol. Mit diesem Wissen kann der Einfluss der räumlich nahen Datenpunkte erhöht werden, wenn der zu vorhersagende Punkt am Nord-bzw. Südpol ist.

Zur Vorhersage können jedoch nur die Datenpunkte verwendet werden, die bereits durchlaufen wurden. Für den ersten Datenpunkt kann deswegen keine Vorhersage gegeben werden. Doch bereits für den zweiten steht der erste Datenpunkt als Informationsquelle zur Verfügung. Für jeden n-ten Punkt gibt es somit genau (n-1) mögliche Informationsquellen. Der Grund für diese Einschränkung ist die Bedingung der Umkehrbarkeit des Kompressionsverfahrens. Die Originaldaten, welche komprimiert werden, müssen schließlich durch die Dekompression wieder zurückgewonnen werden. Deswegen können für die Kompression nur Informationen verwendet werden, welche auch bei der Dekompression zur Verfügung stehen.

Die Kompression entsteht dadurch, dass nur noch die Differenz (das sogenannte Residuum) zwischen der Vorhersage und dem wahren Wert gespeichert werden muss. Hierfür berechnet der Algorithmus zunächst einen Versatz, um diesen auf beide Werte zu addieren. Das hat den Vorteil,

dass das Residuum kleiner ausfällt als es bei der direkten Differenzberechnung wäre und somit ein höherer Kompressionsfaktor erreicht wird.

Danach wird das Residuum in drei Datenströme aufgeteilt. Die ersten beiden Datenströme werden mit dem Burrow-Wheeler-Verfahren umsortiert. Dies erhöht die Wahrscheinlichkeit, dass zwei aufeinanderfolgende Zahlen gleich sind. Danach werden diese Daten mit einer arithmetischen Kodierung versehen, was die finale Kompression darstellt. Der dritte Datenstrom wird schließlich unkomprimiert

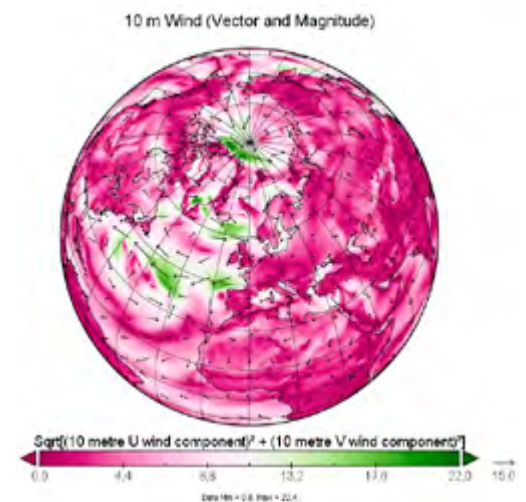


Abbildung 2: Durchschnittliche globale Windgeschwindigkeit im April 2020 auf 10m Höhe (CAM5 Wettermodell).

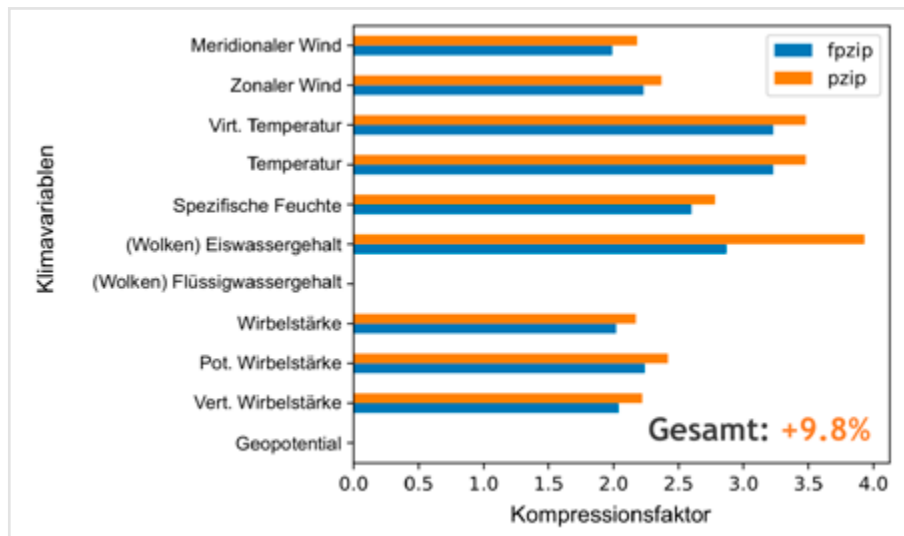


Abbildung 3: Im Durchschnitt erreicht das am SCC und IMK entwickelte pzip Verfahren einen 10% höheren Kompressionsfaktor als der aktuelle Stand der Technik (fpzip) [1].

gespeichert, weil eine verlustfreie Komprimierung mit aktuellen Methoden nicht möglich ist, da er zu sehr einem natürlichen Rauschen ähnelt. Die Originaldaten können nun wiederhergestellt werden, wenn die Traversierungs- und Vorhersagemethode sowie das Residuum bekannt sind.

### Beiträge zur Wissenschaft

Im Rahmen der Arbeit wurden Analysemethoden zur Identifikation von redundanten Informationen in Klimadaten entwickelt, ein Vergleich von verlustfreien Kompressionsverfahren basierend auf Kompressionsfaktor und Durchsatz durchgeführt, ein

neues Datenkodierungsverfahren für vorhersagebasierende Kompressionsverfahren entwickelt, ein neues verlustbehaftetes Kompressionsverfahren für Klima-Indizes entwickelt, das Konzept von Informationsräumen entwickelt und eine neue Datenkodierungsmethode ausgearbeitet.

Der aktuelle Algorithmus erreicht den höchsten Kompressionsfaktor innerhalb der getesteten Verfahren. Der Engpass ist derzeit noch die niedrige Durchsatzrate. Das hängt mit der Implementierung der Kodierungsschritte zusammen. Das Ziel der kommenden Monate ist es, diesen Engpass zu entfernen.

Ich hoffe dieser Artikel konnte zumindest etwas Neugierde zum Thema Wetter- und Klimawissenschaften wecken. Die Wettervorhersage, welche in den Nachrichten vorgetragen nur knapp 60 Sekunden dauert, wird seit dem 20. Jahrhundert stetig weiterentwickelt und verbessert – wie auch die Klimaprojektionen. Auch wenn die Vorhersage einmal falsch sein sollte und wir wider Erwarten im Regen stehen, so ist die Erstellung der Vorhersage doch faszinierend.

### Open Science und Veröffentlichungen

Alle im Rahmen dieser Arbeit entwickelten Algorithmen und Applikationen sowie benutzten Daten, sind auf [www.github.com/ucyo](http://www.github.com/ucyo) unter der GPLv3 Lizenz veröffentlicht und stehen zur freien Verfügung. Eine Sammlung der wissenschaftlichen Veröffentlichungen ist unter [s.kit.edu/ucyo](http://s.kit.edu/ucyo) erreichbar.

[1] publikationen.bibliothek.kit.edu/1000105055

### Danksagung

An dieser Stelle möchte ich mich bei Peter Braesicke, Tobias Kerzenmacher und Jörg Meyer für Ihre Ratschläge und Mitarbeit an der Erstellung dieses Artikels danken.

#### Kompressionsfaktor

Der Kompressionsfaktor setzt die Dateigröße vor und nach der Kompression in Relation zueinander. Hierbei wird die ursprüngliche Dateigröße durch die komprimierte Dateigröße geteilt:

$$\text{Kompressionsfaktor} = \frac{\text{Dateigröße vor der Kompression [Bytes]}}{\text{Dateigröße nach der Kompression [Bytes]}}$$

#### Durchsatz

Der Durchsatz ist die Zeit, welche benötigt wird, um eine Größeneinheit an Daten (üblicherweise Bytes) zu komprimieren bzw. dekomprimieren.

$$\text{Durchsatz} = \frac{\text{Dateigröße [Bytes]}}{\text{(De) Kompressionszeit [s]}}$$

#### Why sometimes less is more – even in science!

In a collaboration, the SCC and the Institute of Meteorology and Climate Research (IMK) have developed new compression algorithms. On average, the algorithms reduce the storage requirements of structured climate data by a compression factor of 2.6 without losing any information. This has implications for all our lives: With limited storage resources, scientists can perform weather forecasts and climate projections with higher accuracy to reliably support a country's climate policy. Data centres can reduce their CO<sub>2</sub> emissions, because a smaller number of simulations are necessary, if more data can be stored for each simulation. Bandwidth utilization is improved because the amount of data to be transferred is reduced. The lifetime of hard disks is extended by reducing their utilization; and scientific simulations can potentially be calculated faster because more data can be stored in main memory.



## ‚Episteme in Bewegung‘ – Veränderlichem Wissen auf der Spur

Der Sonderforschungsbereich 980 ‚Episteme in Bewegung‘ untersucht seit 2012 Prozesse des Wissenswandels in europäischen und nicht-europäischen Kulturen vom 3. Jahrtausend vor Christus bis etwa 1750 nach Christus. Seit 2016 unterstützt das SCC durch seine Expertise im modernen Forschungsdatenmanagement beim Sammeln digitaler Indizien für bisher ungelöste Fragen.

*Danah Tonne, Germaine Götzelmann*

**Welche guten Wünsche für das Jenseits wurden rund 2000 Jahre vor Christus auf dem Sarg des Gaufürsten Mesehti in Assiut aufgebracht, waren aber möglicherweise zur selben Zeit in Saqqara oder Theben völlig unbekannt?**

**Welche händische Seitengestaltung begünstigte das Lehren und Lernen von Aristoteles‘ Logik bis zur Erfindung des Buchdrucks, und wie kann anhand von Abschreibfehlern die Verbreitung der aristotelischen Schriften zwischen Konstantinopel, dem Berg Sinai und Italien erforscht werden?**

**Wie wurde der adeligen Jugend des 17. Jahrhunderts vermittelt, „was zu der Sprache Zierlichkeit zu löblicher Unterhaltung der Gesellschaft und rechter Anstalt des Lebens nothwendig ist“?**

So unterschiedlich die Fragen klingen mögen, so haben diese und viele weitere jedoch den gleichen Kern: Wie verändert sich Wissen in über 4000 Jahren Geschichte? Wie wird es gelehrt, gelernt und in das eigene Leben integriert?

Wissensgeschichtlich wird so ein breites Spektrum von Gegenständen – von der altägyptischen Pyramideninschrift über mittelalterliche Aristotelesabschriften bis zum frühneuzeitlichen Sprachlehrbuch – untersucht, um ein gemeinsames Verständnis von Wissensveränderungen entwickeln zu können. Für einen derart breitgefächerten Forschungsansatz bedarf es allerdings immenser disziplinärer Vielfalt sowie interdisziplinärer Offenheit. Aus diesem Grund bilden Wissenschaftlerin-

nen und Wissenschaftler aus zahlreichen Disziplinen, wie der Ägyptologie, Altorientalistik, Arabistik, Judaistik, Neogräzistik, aber auch Philosophie, Geschichts-, Religions-, Kunst- und Literaturwissenschaften, seit 2012 den Sonderforschungsbereich 980 (SFB 980) ‚Episteme in Bewegung‘. Gefördert von der Deutschen Forschungsgemeinschaft (DFG) vereint der SFB geisteswissenschaftliche und informatische Kompetenzen der Freien Universität Berlin (Sprecher-universität), der Humboldt-Universität zu Berlin, des Max-Planck-Instituts für Wissenschaftsgeschichte und seit 2016 auch des Karlsruher Instituts für Technologie (KIT) sowie der Technischen Universität Darmstadt in einem interdisziplinären Forschungsverbund. Organisiert in 19 Teilprojekten mit insgesamt 55 Mitarbeiterinnen und Mitarbeitern ist dieser SFB daher einer der umfangreichsten Sonderforschungsbereiche auf dem Gebiet der Geistes-, Kultur- und Geschichtswissenschaften.

Dieses seit acht Jahren gelebte Konzept überzeugt nicht nur die Beteiligten, die mit großer Begeisterung Forschungsfragen und -gegenstände aus verschiedenen Perspektiven in den Blick nehmen, sondern auch die Fördergeber. So gab es Ende Mai 2020 Grund zu großer Freude: Die DFG hat die Verlängerung um vier weitere Jahre bewilligt, so dass sich der SFB die (durchaus seltene) maximale Förderdauer von 12 Jahren sichert.

Das SCC ist mit Forschung und Entwicklung im Rahmen des Informationsinfrastrukturprojektes betraut, um gemeinsam mit den Fachwissenschaftlerinnen und

Fachwissenschaftlern und mit Hilfe digitaler Methoden neue Erkenntnisse zu erzielen. In den nächsten vier Jahren liegen die Schwerpunkte auf den folgenden Themenbereichen:

1. **Episteme-Repository:** Von der Hieroglyphenabschrift bis zum Buchdigitalisat werden die Daten und Metadaten aller Teilprojekte im zentralen Repository, basierend auf dem Repositoryframework KIT Data Manager, abgelegt und so mit digitalen Werkzeugen auswertbar.
2. **Vokabulardienste:** Zahlreiche Effekte verhindern allerdings eine automatische Auswertung der Forschungsdaten – beispielsweise zwei Begriffe, die eigentlich dasselbe Phänomen beschreiben (vergleiche Abbildung 2). Mit der Entwicklung einer Dienstlandschaft wird den Fachwissenschaftlerinnen und Fachwissenschaftlern die Möglichkeit geboten, eigene Vokabulare zu erstellen, existierende Vokabulare nachzunutzen sowie mehrere Vokabulare zu kombinieren und damit eine, potentiell auch disziplinübergreifende, Vergleichbarkeit zu erreichen.
3. **Datenanalyse:** Wie unterscheiden sich ägyptische Sargsprüche? Wer hat welches griechische Wort von wem beschrieben? Zum Nachweis von Ähnlichkeiten zwischen Datensätzen werden digitale Werkzeuge benö-

<sup>1</sup> Vgl. Titelblatt zu Johann Hermann Widerhold: Tresor en trois langues, Basel 1679



tigt (z.B. Annotationsumgebungen, Clusteranalyse oder Visualisierung von Merkmalen wie in den Abbildungen 1 und 3 gezeigt) und daher nahtlos in die Infrastruktur eingebunden.

4. **Ergebnispräsentation:** Welche Möglichkeiten gibt es, Verwandtschaft von Texten darzustellen? Und wie können Gemeinsamkeiten und Unterschiede

mit Hilfe von Bildern belegt werden, die man nicht veröffentlichen darf? Dringlich ist daher die Entwicklung von Methoden und Werkzeugen, um die vielfältigen Forschungsergebnisse des SFB öffentlich sichtbar zu machen.

Spannende Gegenstände, vielfältige Herausforderungen und nicht zuletzt gemeinsame Forschung mit unverzichtbaren

Kolleginnen und Kollegen liegen vor uns, so dass sich alle einig sind: Wir freuen uns auf die nächsten vier Jahre mit den ‚bewegten Epistemikern‘!

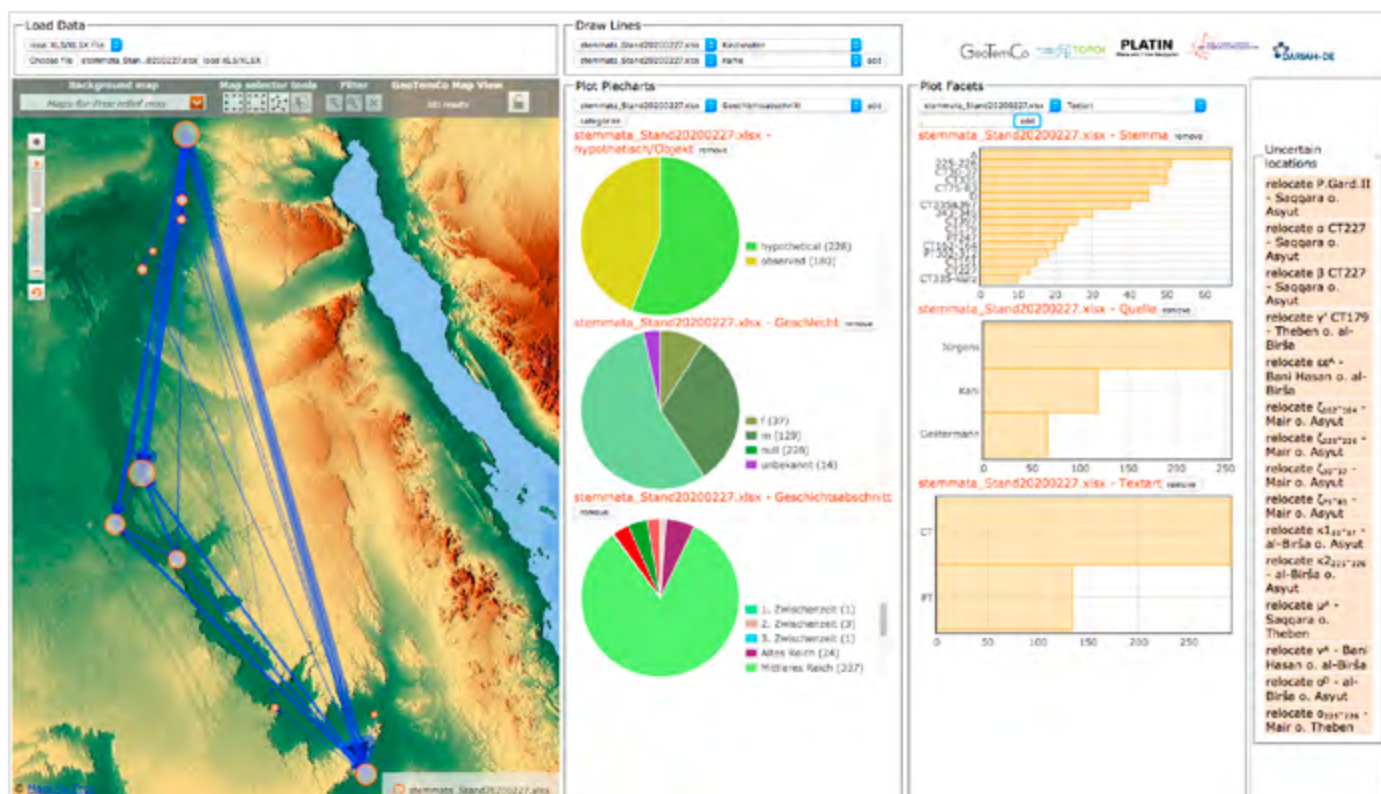


Abbildung 1: Interaktive Visualisierung der Weitergabe ägyptischer Sarg- und Pyramidentexte zur Aufdeckung von Einflussfaktoren wie Ort, Zeit, Geschlecht, soziale Strukturen, uvm. Auf diese Weise wird beispielsweise untersucht, ob ein in Assiut verwendeter Text in Saqqara oder Theben nachgenutzt wurde und welcher zeitliche Abstand dazwischen lag. Grafik wurde mit der Visualisierungssoftware PLATIN (platin.mpiwg-berlin.mpg.de) erstellt.

**Weitere Informationen**

1. Sonderforschungsbereich ‚Episteme in Bewegung‘: [www.sfb-episteme.de/](http://www.sfb-episteme.de/)
2. M. Krewet, P. Hegel, G. Götzelmann, S. Söring, D. Tonne, „Aristoteles auf Reisen: Handschriftenforschung in der digitalen Infrastruktur des SFB 980 ‚Episteme in Bewegung‘“, Seiten 77-87, Forschungsinfrastrukturen in den digitalen Geisteswissenschaften: Wie verändern digitale Infrastrukturen die Praxis der Geisteswissenschaften?, 2019, URN: urn:nbn:de:hebis:30:3-519405
3. G. Götzelmann, P. Hegel, M. Krewet, S. Söring, D. Tonne, „Aspekte digitaler Infrastrukturen. Herausforderungen und Perspektiven von Forschungsdaten im Sonderforschungsbereich 980 ‚Episteme in Bewegung‘“, Seiten 324-331, Bibliothek – Forschung und Praxis 43.2 (2019), DOI: 10.1515/bfp-2019-2055

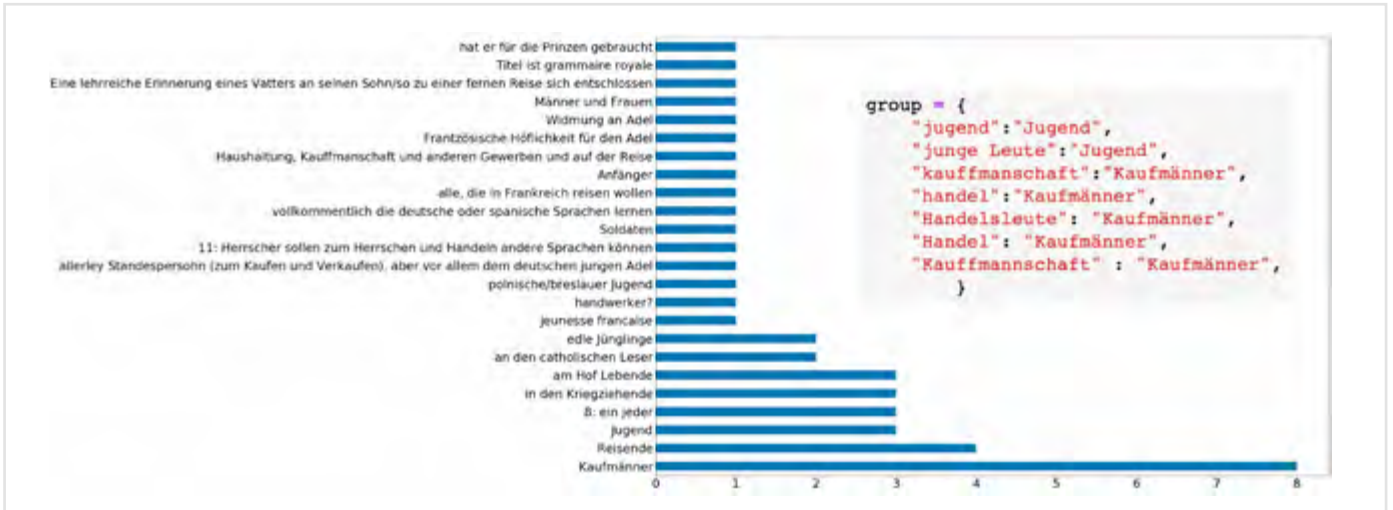


Abbildung 2: Manuell erfasste Zielgruppen von frühneuzeitlichen Sprachlernbüchern als Anwendungsfall für Vokabularien. Nur durch eine Kategorisierung von Begriffen wie ‚Adel‘ und ‚Jugend‘ wird der gesamte Buchbestand mit digitalen Werkzeugen auswertbar.

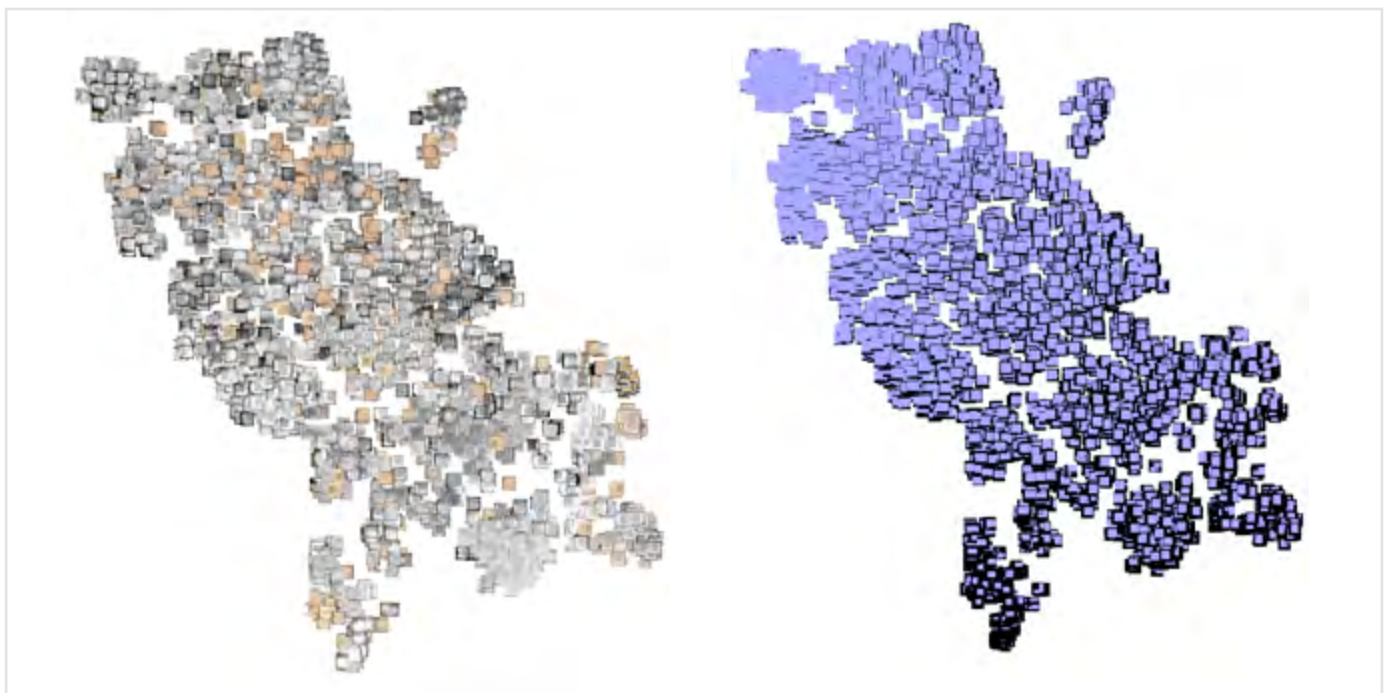


Abbildung 3: Visualisierung der überlieferten Exemplare von Aristoteles' Schrift de interpretatione (links: Originalseite, rechts: erkannte Textregionen, geordnet nach Verhältnis von Text zu Seite). Auf diese Weise kann die Seitengestaltung einzelner Manuskripte/Manuskriptgruppen verglichen werden, um beispielsweise Anzahl und Häufungen von Kommentaren und damit die Lehr-/Lernpraxis aufzudecken.

### 'Episteme in Motion' – Tracking Mutable Knowledge

Since 2012, the Collaborative Research Centre 'Episteme in Motion' has been investigating processes of knowledge change in European and non-European cultures from the 3rd millennium BC to about 1750 AD. At the end of May 2020 there was reason for great joy: The German Research Foundation has approved the extension for another four years, thus securing the maximum funding period of 12 years for the CRC. The SCC is conducting research and development within the information infrastructure project of the CRC in the areas of research data repositories, vocabulary services, as well as data analysis, exploitation, and visualization.

## Peer Review 2.0

In fast allen bahnbrechenden Forschungserfolgen der letzten Jahre spielen Computersimulationen und -analysen eine zentrale Rolle. Die hier vorherrschende Schnelligkeit äußert sich jedoch in einer stark wachsenden Anzahl zu begutachtender wissenschaftlicher Publikationen. Die Begutachtung für Fachzeitschriften bestimmt daher immer stärker den Arbeitsalltag der Forschenden, so dass das Peer-Review-Verfahren in seiner derzeitigen Form bald an seine Kapazitätsgrenzen stoßen wird. Die Entwicklung eines nachhaltigen Peer-Review-Verfahrens auf Basis etablierter Prozesse von Open-Source-Projekten, wie in diesem Beitrag geschildert, trägt dazu bei, dass der Aufwand für Begutachtungen minimiert werden kann, bei gleichzeitiger Verbesserung der Überprüfbarkeit und Verfügbarkeit von Software für und in der Forschung.

*Eileen Kühn, Hartwig Anzt*

Forschende stehen heutzutage vor ganz besonderen Herausforderungen, denn in vielen Bereichen werden konventionelle Laborexperimente und Messreihen durch komplexe Computersimulationen und Analysen großer Datenmengen ergänzt oder sogar fast vollständig verdrängt. Zusammen mit Fachkundigen der Domänen- und Datenwissenschaften, der Mathematik, der Informatik und des maschinellen Lernens werden Algorithmen und Analysewerkzeuge entwickelt und erfolgreich für die Forschung eingesetzt.

Die Kombination interdisziplinärer Forschung und neuer, agiler Forschungsmethoden in den computergestützten Disziplinen wie z.B. Klimamodellierung, Wetter- und Erdbebenvorhersage oder Astrophysik-Anwendungen ermöglicht ein bisher nicht dagewesenes Innovationspotential. So spielen in fast allen bahnbrechenden Forschungserfolgen der letzten Jahre Computersimulationen und -analysen eine zentrale Rolle. Insbesondere die Schnelligkeit dieser zukunftssträchtigen Disziplinen bedingt jedoch ein stark umkämpftes Forschungsfeld, was sich in einer extrem steigenden Anzahl wissenschaftlicher Publikationen im Bereich der Algorithmik und des wissenschaftlichen Rechnens äußert. Eine Vielzahl der Publikationen basiert auf prototypischen Softwareentwicklungen, deren vorrangiges Ziel die schnelle Veröffentlichung vorläufiger Ergebnisse darstellt. Diese Arbeitsweise optimiert zwar den Nutzen für den Einzelnen, der Nutzen für die Wissenschaft ist jedoch fragwürdig:

Prototypische Softwareentwicklungen erschweren nicht nur die Adaption und Reproduktion durch Teammitglieder oder verwandte Fachbereiche, sondern erfordern meist einen erheblichen Aufwand, um die vorgestellten Lösungen langfristig durch hochwertige Neuimplementierungen für die Wissenschaftsgemeinschaft und darüber hinaus nutzbar zu machen.

### Die Bedeutung des Peer-Review-Verfahrens für die Wissenschaft

Bereits über Jahrhunderte wird die Qualität wissenschaftlicher Veröffentlichungen durch das Peer-Review-Verfahren nach Henry Oldenburg (s. Kasten) sichergestellt. Gerade dieser Erfolg begründet das heutige Maß für wissenschaftliche Exzellenz: die Anzahl an Veröffentlichungen und deren Zitierungen (der sogenannte H-Index) als zentrale Leistungskennzahlen für die akademische Reputation. Die zunehmende Relevanz solcher Kennzahlen im Wettbewerb um Forschungsmittel und als Basis für eine wissenschaftliche Karriere erzeugt einen stetig wachsenden Druck bei Forschenden, Projektkonsortien und Forschungseinrich-

tungen, eine hohe Anzahl an Fachbeiträgen einzureichen. Das wiederum zieht einen Anstieg an Gutachtertätigkeiten für Wissenschaftler und Wissenschaftlerinnen nach sich.

Fundierte wissenschaftliche Gutachten zu erstellen erfordert bereits heute einen signifikanten Arbeitsaufwand. Dabei wird die Review-Tätigkeit für Forschende meist weder finanziell ausgeglichen, noch trägt sie zur Reputation des Reviewers bei. Dem steigenden Druck zur Veröffentlichung steht somit kein vergleichbarer Anreiz zur gewissenhaften Begutachtung dieser Veröffentlichung gegenüber. Wenn auch in Zukunft der H-Index unverändert als

### Das akademische Peer-Review-Verfahren nach Henry Oldenburg

Das akademische Peer-Review-Verfahren wurde von Henry Oldenburg bereits im 17. Jahrhundert eingeführt. Auch heute noch wird dieses Verfahren nahezu unverändert bei der Veröffentlichung von Forschungsergebnissen eingesetzt und hat sich als wesentliches Werkzeug der Qualitätskontrolle und -sicherung durchgesetzt:

- 1) Forschende reichen wissenschaftliche Artikel zur Präsentation neuer Forschungsergebnisse in Fachjournalen oder -konferenzen ein,
- 2) ein Editor ernennt unabhängige Gutachter des gleichen Fachgebiets, sogenannte Peers,
- 3) die Begutachtenden bewerten den Gehalt, die Qualität, die Neuartigkeit und die Korrektheit der Forschung und
- 4) basierend auf diesen Peer Reviews wird die Einreichung veröffentlicht, abgelehnt oder eine Überarbeitung des Artikels empfohlen.

zentrale Kennzahl für die akademische Reputation bestehen bleibt, wird das Peer-Review-Verfahren durch die wachsende Review-Tätigkeit bald an seine Kapazitätsgrenzen stoßen. Als zentrale Säule des wis-



wissenschaftlichen Arbeitens gefährdet dies nicht nur den Prozess des Peer-Review-Verfahrens selbst, sondern den gesamten Wissenschaftsbetrieb.

Gleichzeitig wird als Teil der Forschungsstrategie insbesondere in den computergestützten Wissenschaften der Prozess einer guten und nachhaltigen Softwareentwicklung angestrebt: Wo Forschende ursprünglich prototypisch Funktionen schnell implementieren konnten, sehen sie sich nun mit neuen Herausforderungen konfrontiert, die zusätzlich zu den bestehenden Review-Tätigkeiten einer schnellen Publikation ihrer Forschungsergebnisse entgegenlaufen.

**Die Bedeutung nachhaltiger Softwareentwicklung in den computergestützten Wissenschaften**

Parallel zum Peer-Review-Verfahren hat sich in der Softwareentwicklung ein Bewusstsein für gegenseitige Begutachtung, Überprüfbarkeit und Verfügbarkeit entwickelt – nicht zuletzt durch die Mitwirkung von Anwendern und Anwenderinnen aus den unterschiedlichen Wissenschaftsdisziplinen. Für die Softwareentwicklung und insbesondere bei Open-Source-Projekten werden kollaborative Qualitätssicherungskonzepte verwendet, bei denen alle Änderungen automatisiert und durch die Community auf Korrektheit, Dokumentation und Nachhaltigkeit überprüft werden (s. Kasten). Solange wissenschaftliche Software nicht diesen Prozess durchläuft, sondern im Prototypstatus entsteht und dauerhaft verbleibt, kann das Innovationspotential in computergestützten Disziplinen nicht voll ausgenutzt werden. Leider genügen die meisten Softwareentwicklungen nicht den Kriterien nachhaltiger Open-Source-Projekte. Teil des Problems ist die mangelnde Bekanntheit der Rolle des Research Software Engineers<sup>1</sup> (RSE) im Wissenschaftsbetrieb sowie die fehlende Wertschätzung *nachhaltiger Softwareentwicklung* für die Wissenschaft. Insbesondere im Vergleich zu begutachteten Veröffentlichungen verhilft die nachhaltige Softwareentwicklung einem einzelnen

Wissenschaftler kaum zu akademischer Anerkennung. Es gibt eine massive Diskrepanz zwischen dem, was Forschende zur Erhaltung und Steigerung ihrer Reputation tun müssen – publizieren auf Basis von Prototypsoftware – und dem, was der wissenschaftlichen Gemeinschaft langfristig nutzt: reproduzierbare Forschung durch hochwertige, nachhaltige Softwareentwicklung.

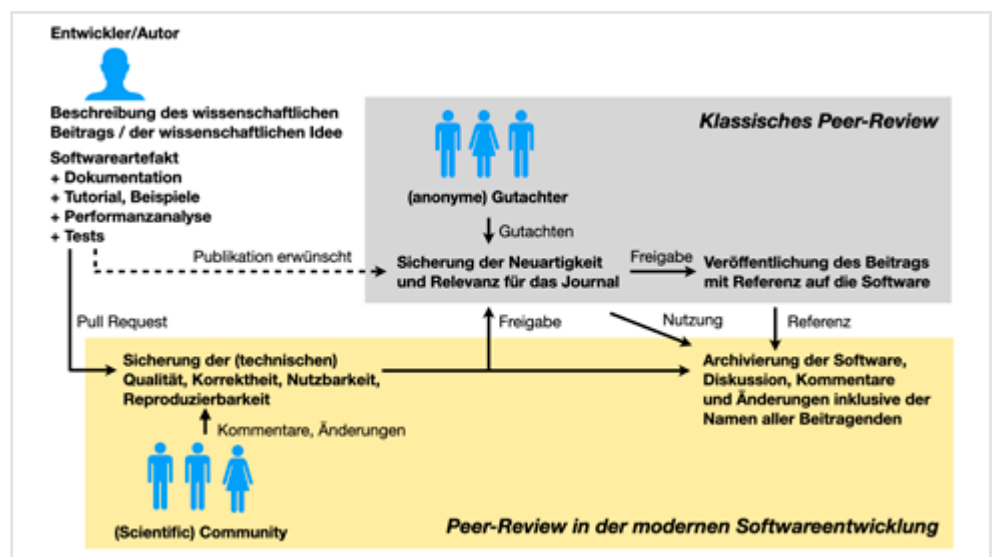
Diese Parallelentwicklung des Software-Review-Verfahrens ermöglicht es, dass Open-Source-Projekte eine der zentralen Komponenten für das Innovationspotential und die Nachhaltigkeit der computergestützten Wissenschaften bilden. Für das Peer-Review in den computergestützten Wissenschaften kann man diese Parallelstrukturen sowohl als ein weiteres System als auch redundant interpretieren. Wir sehen es als einzigartige Chance, die Begutachtung, Überprüfbarkeit und Verfügbarkeit von Software für und in der Forschung als essentielles Standbein zu etablieren.

**Ein nachhaltiges Peer-Review-Verfahren für die computergestützten Wissenschaften**

Können auf Basis des aktuellen Peer-Review-Verfahrens, dass bereits an seine

Grenzen von Qualität und Quantität stößt, noch weitere und verschärfte Anforderungen um- und durchgesetzt werden? Wie viel Zeit können Forschende der computergestützten Wissenschaften in Softwareentwicklungen investieren, ohne dass dies entsprechend honoriert wird? Können freiwillige Kollaborationen, deren Ziel die Entwicklung nachhaltig nutzbarer Software ist, sich einer stetig wachsenden Zahl von Prototypen annehmen? Unter diesen Gesichtspunkten scheint es konsequent, das klassische Peer-Review-Verfahren und die Bedeutung der zentralen Leistungskennzahlen in Frage zu stellen. Da sich die klassische Publikation und die daraus abgeleiteten Metriken mit all ihren Vor- und Nachteilen fest im Wissenschaftssystem etabliert haben, ist es jedoch eine Herausforderung, das bewährte System an die modernen Gegebenheiten anzupassen. Ein radikaler Umstieg auf neue Metriken, wie es etwa die San Francisco Declaration on Research Assessment (DORA) fordert<sup>2</sup>, erscheint deshalb kurzfristig unrealistisch und kann eher als mittel- bis langfristige Strategie betrachtet werden.

Wir wollen stattdessen den Prozess der wissenschaftlichen Begutachtung von Forschungsergebnissen an die Arbeitsweise der modernen, computergestützten Wissenschaften anpassen, um 1) den



Das nachhaltige Peer-Review-Konzept für die computergestützten Wissenschaften erweitert das Community-basierte Software-Review-Verfahren um anonyme Gutachten zur Evaluation des wissenschaftlichen Innovationsgrads.

<sup>1</sup> Die RSEs kombinieren zwei wertvollen Fähigkeiten: Fachwissen in der Programmierung und ein gutes Verständnis komplexer Forschung. (Quelle: software.ac.uk/research-software-engineers)

<sup>2</sup> sfdora.org



## Workflow zur Überprüfung von Softwarebeiträgen in Open-Source-Projekten

Etablierte Open-Source-Softwareprojekte, wie z.B. TensorFlow, Pandas, PyTorch, Trilinos, PetSC oder OpenFOAM setzen sorgfältige Software-Review-Verfahren ein, bei denen jeder Softwarebeitrag (Pull Request) von anderen Entwicklern in Bezug auf Korrektheit, Nutzen, Kompatibilität, Qualität und Vollständigkeit begutachtet, kommentiert und bezüglich seiner Integration bewertet wird. In Verbindung mit der freien Interaktion mit der Community führt dies in der Regel zu erstklassig dokumentierten und qualitativ hochwertigen Softwarebeiträgen.

Zentrale Komponente um solch ein Software-Review-Verfahren zu etablieren, ist ein modernes Versionierungssystem wie z.B. Git oder Mercurial. Moderne Versionierungssysteme unterstützen nicht nur bei der gemeinsamen, verteilten Entwicklung, Protokollierung, Archivierung und Rückverfolgbarkeit, sondern ermöglichen außerdem die direkte Integration zusätzlicher Funktionen, wie automatisierte Tests oder Kontrollen zur Testabdeckung und gewährleisten so die Qualität und Nachhaltigkeit der Softwareentwicklung.

Begutachtungsaufwand für Forschende zu minimieren, 2) den durch die Gutachter geleisteten Aufwand in und über die Forschungsgemeinschaft hinaus anzuerkennen und in der akademischen Reputation zu reflektieren, 3) die Produktivität des Einzelnen und das Innovationspotential der Gemeinschaft durch nachhaltige Softwareentwicklung zu steigern sowie 4) eine vollständige Reproduzierbarkeit und Rückverfolgbarkeit der präsentierten Ergebnisse zu ermöglichen.

Unsere zentrale Idee zur Modernisierung des Peer-Reviews besteht darin, das Software-Review-Verfahren als Grundlage für ein nachhaltiges Peer-Review-Konzept in den computergestützten Wissenschaften zu nutzen und um ein klassisches Peer-Review-Verfahren zu erweitern. Im Kern basiert das nachhaltige Peer-Review-Konzept also auf den bestehenden Methoden moderner Softwareentwicklung rund um verteilte, kollaborative Versionierungssysteme. Da Versionierungssysteme alle Beiträge und Diskussionen archivieren, ist die vollständige Reproduzierbarkeit und Rückverfolgbarkeit garantiert. Automatisierung von Tests und die stetige Validierung machen eine Überprüfung auch von komplexen Ergebnissen sowohl effizient als auch nachvollziehbar. Außerdem erhalten nicht nur die einreichenden Autorinnen und Autoren, sondern auch alle beitragenden Gutachterinnen und Gutachter nachhaltig Aufmerksamkeit für ihre Beiträge, Empfehlungen und Kritiken, die somit eine ausgezeichnete Grundlage für die Publikation in Konferenzen und Journalen bilden. Der stetige, offene und hochwertige Austausch zwischen Wis-

senschaftlern und Wissenschaftlerinnen, die ihre Forschung auf ein gemeinsames Softwareprojekt stützen, ist somit auch ein nachhaltiges Peer-Review, das dem klassischen Peer-Review-Verfahren in nichts nachsteht.

Wir schlagen deshalb vor, in publikationsrelevanten Fällen den nachhaltigen Peer-Review-Prozess zu durchlaufen, also den Software-Review-Prozess um ein zusätzliches Gutachten zu erweitern, das evaluiert, ob ein Softwarebeitrag veröffentlichungswürdig ist.

Abbildung 1 illustriert die Komponenten und Beteiligten des nachhaltigen Peer-Review-Konzeptes: Zusätzlich zu dem Software-Gutachten des Pull Requests kann ein Autor, eine Autorin einen Softwarebeitrag im Versionierungssystem als Beitrag für eine bestimmte Konferenz oder ein Journal markieren, um den Editor darauf hinzuweisen, dass ein weiteres, unabhängiges und anonymes Gutachten erwünscht ist. Der Editor beauftragt daraufhin anonyme Gutachten unabhängiger Peers, die den Innovationsgrad des Beitrags beurteilen. Sowohl die technische Qualität als auch die Reproduzierbarkeit der Ergebnisse kann bereits vor der eigentlichen Begutachtung im Sinne des klassischen Peer-Review-Verfahrens durch eine breite Entwickler-Gemeinschaft sichergestellt und fortwährend problemlos überprüft und nachvollzogen werden. Dies minimiert den Bearbeitungsaufwand für die externen Begutachtenden erheblich, da sie jederzeit vollen Zugriff auf die Community-Reviews im Versionierungssystem haben und ihre Hauptaufgabe damit in der

Beurteilung der Relevanz und Neuartigkeit des umgesetzten Konzepts besteht.

Auf Basis der Erfahrungen mit dem Journal on Open Source Software, das zunehmend Akzeptanz in den computergestützten Wissenschaften findet, sind wir davon überzeugt, dass ein solches Community-basiertes Peer-Review-Konzept nicht nur den derzeitigen Druck auf das traditionelle Peer-Review-Verfahren reduziert, sondern auch die Qualität und Dokumentation von Forschungssoftware nachhaltig steigert – ein zentrales Ziel aller nationalen und internationalen Initiativen zur Förderung von Forschungssoftware.

Zur Evaluierung der Umsetzbarkeit und Akzeptanz des Konzeptes werden wir beim Workshop on Scalable Data Analytics in Scientific Computing<sup>1</sup> (SDASC) bei der ISC 2021 erstmals Beiträge dieser Form erlauben.

<sup>1</sup> Der vollständige Aufsatz ist zu finden unter: [onlinelibrary.wiley.com/doi/abs/10.1002/pamm.201900490](https://onlinelibrary.wiley.com/doi/abs/10.1002/pamm.201900490)

### Peer Review 2.0

Computer simulations and analyses play a key role in almost all groundbreaking research successes of recent years. However, the fast pace of this scientific field is reflected in a rapidly growing number of scientific publications that need to be re-viewed. Therefore, peer reviews for scientific journals are increasingly determining the everyday work of researchers, so that the peer review process in its current form will soon reach its limits. The development of a sustainable peer review process based on established processes of open source projects, as described in this article, will help minimising the effort for peer reviews while improving the verification and availability of software for and in research.

# Helmholtz Inkubator Pilotprojekt Uncertainty Quantification

Das Helmholtz Inkubator Pilotprojekt Uncertainty Quantification – From Data to Reliable Knowledge hat zum Ziel, mit Methoden der Unsicherheitsquantifizierung zuverlässigere Erkenntnisgewinne aus Daten zu ermöglichen. Eine Besonderheit des Projekts besteht darin, Forschende unterschiedlichster Wissenschaftsbereiche zu vernetzen, um breit anwendbare mathematische und numerische Werkzeuge und Methoden im Bereich der Uncertainty Quantification zu entwickeln und auf die einzelnen Fachbereiche anzuwenden.

*Jonas Kusch, Samuel Braun*

## Was ist Uncertainty Quantification?

Unsere Welt, und insbesondere deren Erforschung, ist geprägt von Unsicherheiten. Anwendungen, in denen die Effekte von Unsicherheiten die gewonnenen wissenschaftlichen Erkenntnisse maßgeblich beeinflussen, sind zum Beispiel mathematische Modelle zur Vorhersage der Entwicklung der Corona-Pandemie. Hier sind die exakte Anzahl der Infizierten sowie Modellierungsparameter wie die Reproduktionszahl nicht bekannt. Diese Unsicherheiten müssen in Simulationen berücksichtigt werden, um eine geeignete Aussage zur Pandemie-Entwicklung zu treffen. Weitere Beispiele sind die Wettervorhersage, in der die anfängliche Temperaturverteilung unsicher ist, oder Anwendungen im Bereich der künstlichen Intelligenz, bei denen die Auswirkung von fehlerbehafteten Trainingsdaten in getroffenen Entscheidungen mit berücksichtigt werden sollten.

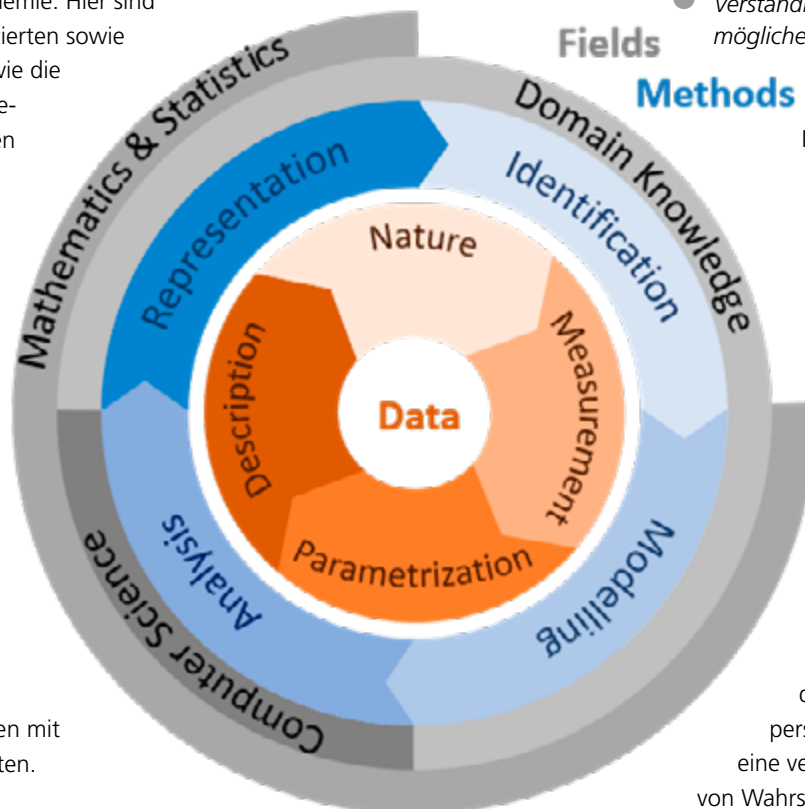
Die Unsicherheitsquantifizierung (engl. Uncertainty Quantification) ist ein interdisziplinäres Forschungsgebiet, das sich zum Ziel setzt, diese Unsicherheiten zu beschreiben und die Effekte dieser Unsicherheiten auf die Vorhersagegenauigkeit von Modellen abzuschätzen. Hierzu bietet die Unsicherheitsquantifizierung verschiedene mathematische Methoden.

## Projektziele und das interdisziplinäre Team

Das Helmholtz-Projekt Uncertainty Quantification (UQ) hat zum Ziel, mit Methoden der Unsicherheitsquantifizierung zuverlässigere Erkenntnisgewinne aus Daten zu ermöglichen, indem breit anwendbare

Information untereinander sowie die Helmholtz-Datenwissenschaftler und externe Hochschulpartner aus Mathematik und Ökonometrie.

*Abbildung 1: Quellen für Unsicherheiten (braun) und Methoden zur Quantifizierung von Unsicherheiten (blau). Das Zusammenspiel verschiedener Felder führt zu einem besseren Verständnis von Unsicherheiten und deren möglichen Auswirkungen.*



Mit Hilfe der entwickelten Werkzeuge und methodischen Ansätze werden Anwendungswissenschaftlerinnen verlässliches Wissen aus Unsicherheit extrahieren können, wo traditionelle Ansätze und Software an den schnell wachsenden Datenmengen und deren hoher Variabilität scheitern. Gerade bei Themen von öffentlichem Interesse, wie Klimawandel, Energieversorgung und personalisierter Medizin, ist eine verständliche Vermittlung von Wahrscheinlichkeiten und Risiken unerlässlich. Unsicherheiten sollen besser verstanden werden, und die Datenkompetenz unter Helmholtz-Forschenden und über die Grenzen der Gemeinschaft hinaus soll gestärkt werden.

Werkzeuge und Methoden im Bereich der Uncertainty Quantification anhand der Anwendungsfälle entwickelt werden (Abbildung 1). Das Projekt vernetzt angewandt Forschende aus den vier Helmholtz-Forschungsbereichen Erde & Umwelt, Energie, Gesundheit und Infor-

Das SCC trägt an unterschiedlichen Stellen zum Gelingen des Projekts bei. Zum einen werden die unterschiedlichen

Anwendungsbereiche mit mathematischer und methodischer Kompetenz unterstützt, zum andern ist das KIT in der Projektkoordination involviert.

### **Agiles Arbeiten in der Wissenschaft?!**

Ebenso wie in Wirtschaft und Industrie ist ein großer Aspekt in der Wissenschaft die Bearbeitung und das Management von Projekten, welches maßgeblich über Erfolg oder Misserfolg entscheidet. Insbesondere in der Wirtschaft spielt daher die Untersuchung verschiedener Projektmanagementstrategien eine große

### **Helmholtz Incubator Pilot Project Uncertainty Quantification**

The Helmholtz Incubator Pilot Project Uncertainty Quantification - From Data to Reliable Knowledge aims to enable more reliable insights from data using methods of uncertainty quantification. A special feature of the project is the networking of researchers from different scientific fields in order to develop broadly applicable mathematical and numerical tools and methods of Uncertainty Quantification and apply them to the individual disciplines.

Rolle. Hierbei stellt sich heraus, dass sogenannte agile Methoden gut geeignet sind, um Kreativität und Produktivität zu fördern. Wir als Wissenschaftlerinnen und Wissenschaftler wollen von dieser Erkenntnis profitieren und in der Wirtschaft erprobte Projektmanagementmethoden im wissenschaftlichen Kontext anwenden. Das hier gewählte Verfahren ist Scrum, welches Projekte in ein- bis vierwöchigen sogenannten Sprints bearbeitet. Wichtige Rollen im Scrum-Prozess sind das Entwicklerteam, das während des Sprints das gewünschte Produkt weiterentwickelt, der Scrum Master, der versucht, Hindernisse zu beseitigen und auf die Einhaltung des Scrum-Prozesses zu achten, sowie der Product Owner, der dafür sorgt, dass die Entwicklung zum festgelegten Ziel führt.

Zum besseren Verständnis des Scrum-Prozesses wurden die Mitglieder des Projektes UQ in mehreren Scrum-Schulungen mit der Methode vertraut gemacht. Ein erster Sprint erfolgte zur Entwicklung der Projektwebsite. Hierbei stießen viele der von SCRUM verfolgten Ideen und Konzepte auf Anerkennung. Einige Aspekte, die den Fokus zu stark auf die Arbeit in Unternehmen legen, müssen jedoch noch an das wissenschaftliche Arbeiten angepasst werden.

### **Wie geht es weiter?**

Das Kickoff Meeting fand im November 2019 in Bielefeld statt. Mittlerweile konnten die meisten der vorgesehenen Promotionsstellen besetzt werden. Wegen der eingeschränkten Reisemöglichkeiten aufgrund der Corona-Pandemie im 2. Quartal dieses Jahres gestalteten sich die Einstellungsprozeduren, insbesondere für ausländische Promovierende, schwieriger als sonst. Ebenfalls Pandemiebedingt wurden das Scrum-Training Ende April und ein erstes Projekttreffen als Onlineformate konzipiert und durchgeführt. Um das im Scrum-Lehrgang Gelernte anzuwenden und zu festigen, wurde in einem einwöchigen Sprint als erstes gemeinsames Trainingsprojekt die Projektwebsite erstellt. Als erstes vorbereitendes wissenschaftliches Projekt steht die Entwicklung eines UQ-Wörterbuchs an, mit dem Ziel, eine gemeinsame Sprache und einheitliche Begriffe zu finden, was die Behandlung und Benennung von Unsicherheiten in den verschiedenen Fachbereichen angeht.

[www.helmholtz-uq.de/](http://www.helmholtz-uq.de/)

## CAMMP goes online

Um die Ansteckungsgefahr mit dem Corona-Virus einzudämmen, hat das Ministerium für Kultus, Jugend und Sport bis zu den Sommerferien alle außerschulischen Veranstaltungen, zu denen auch die CAMMP days des Projekts Computational and Mathematical Modeling Program am KIT gehören, verboten. So sah sich das CAMMP-Team gefordert, ein Onlineprogramm zu etablieren und durchzuführen.

*Maren Hattebuhr, Sarah Schönbrodt*

Wir wissenschaftlichen Mitarbeiterinnen werden nach Hause geschickt, um von dort weiter zu arbeiten. Denn die Corona-Krise erfordert es, nach den erlassenen Verordnungen zu handeln. Das ist auch alles kein Problem: Den Laptop einpacken, noch einen Bildschirm, eine Tastatur und Maus mitnehmen, dann ist der Arbeitsplatz daheim gut eingerichtet.

Für CAMMP<sup>1</sup> und die angebotenen Schüleraktivitäten steht jedoch zunächst erstmal alles Kopf: Das Schulministerium hatte angekündigt, dass alle außerschulischen Veranstaltungen, zu denen auch unsere CAMMP days gehören, bis zu den Sommerferien verboten werden! Es hatten sich schon viele Schulen angemeldet, um insbesondere auch unsere neu entwickelten Lernmaterialien zu testen. Alles musste nun abgesagt werden. Dies ist nicht nur ein herber Rückschlag für CAMMP, sondern auch für uns Doktorandinnen, die auf die Evaluation des Lernmaterials durch Schüler/innen angewiesen sind. Daher war schnell klar: CAMMP goes online.

Zum Glück konnten wir auf die Arbeit aus dem letzten Sommer aufbauen – denn bereits da hatten wir begonnen, sämtliches Lernmaterial digital aufzubereiten und online zur Verfügung zu stellen. Dazu wurden die Arbeitsmaterialien als Jupyter Notebook<sup>2</sup> realisiert. Diese Notebooks sind hervorragend geeignet, um Arbeitsblätter zu entwickeln, bei denen Textbausteine (Aufgabenstellungen, Hintergrundwissen, Lösungshinweise etc.) leicht lesbar und übersichtlich mit Code-Feldern und graphischen Darstellungen verbunden werden (s. Abbildung 1). Die Materialien sind so gestaltet, dass die Lösungen der Schüler/innen im Hintergrund automatisch überprüft werden. Die Lernenden erhalten dann individuelle und auch fehlerspezifische Rückmeldungen zu ihren Lösungseingaben im Code. Dadurch wissen sie einerseits, ob ihre Lösung richtig ist, oder sie bekommen andererseits direkt eine Rückmeldung, in welche Richtung sie ihre Lösung verbessern müssen. Außerdem lassen sich Tipps, Zusatzinformationen oder unterstützende Erklärvideos leicht in Jupyter Notebook verlinken. Aus didaktischer Sicht ist dieses Paket optimal geeignet, um ein selbstbestimmtes Lernen in heterogenen Lerngruppen zu fördern. Darüber hinaus bietet Jupyter Notebook auch technische Vorteile: Es ist eine Open-Source Software, die wir auf einem Server der bwCloud hosten können. Das bietet den großen Vorteil, dass die Nutzer/innen, also Schüler/innen, Lehrkräfte und Studierende, keine Software mehr auf ihren Computern

installieren müssen: Die Bearbeitung der Lernmaterialien erfolgt ausschließlich über einen Webbrowser (Abbildung 2).



*Abbildung 1: Die CAMMP-Lernumgebung nutzt flexibel unterschiedliche Geräte. Beispielsweise kann eine Betreuerin mit Hilfe eines Tablets direkt in die Jupyter Notebooks der Lernenden hineinzeichnen.*

Bleibt nur die Frage: Wie ersetzen wir die persönliche Betreuung der Schüler/innen bei der Bearbeitung der Lernmaterialien? Normalerweise, d. h. wenn Schulklassen zu uns ans KIT kommen, führen zwei Doktorandinnen oder studentische Hilfskräfte (Hiwis) in die Problemstellung ein, beantworten Fragen und machen Mut, falls die „Mathe dann doch mal doof ist“.

Diese persönliche Betreuung findet nun über eine virtuelle Kommunikationsplattform statt. Diese Plattform muss diverse Anforderungen erfüllen: Sie soll sowohl den mündlich als auch den schriftlichen Austausch innerhalb einer Gruppe von etwa



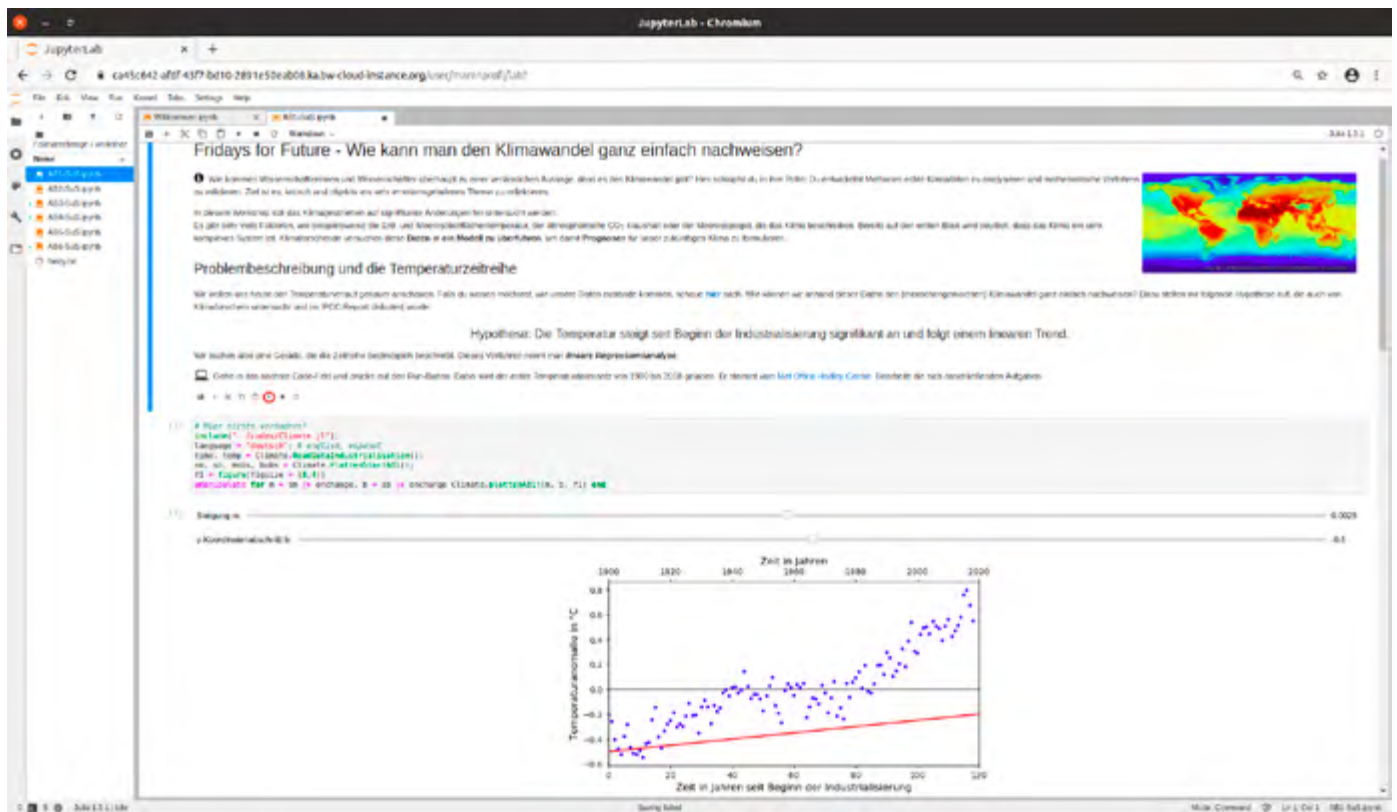


Abbildung 2: CAMMP-Arbeitsmaterialien als Jupyter Notebook realisiert

30 Schüler/innen ermöglichen. Darüber hinaus werden virtuelle Unterräume benötigt, in denen die Schüler/innen in eigenständigen Arbeitsphasen in kleinen Teams diskutieren können. Es wird eine Funktion zum Präsentieren von Vorträgen sowie das Teilen von Bildschirmhalten gefordert. Zu guter Letzt sollte die Plattform im besten Fall ein Whiteboard zur Verfügung stellen, über das alle Mitglieder Ideen austauschen können. All diese Funktionen sollten übersichtlich angeordnet, leicht und intuitiv verwendbar und, ebenso wie Jupyter Notebook, im Webbrowser bedienbar und in der bwCloud hostfähig sein. Die Plattform sollte außerdem unabhängig von der Internetverbindung Einzelner stabil laufen.

Eine Plattform, die all diesen Anforderungen gerecht wird, wurde noch nicht gefunden. Gute Erfahrungen haben wir mit der Chat-

und Videoanwendung Discord<sup>3</sup> gesammelt, die jedoch leider nicht quelloffen ist. Deswegen wird nun die Open-Source-Plattform Mattermost<sup>4</sup> erprobt, was zum jetzigen Stand eine gute Alternative zu sein scheint.

Erfahrungen mit dem neuen Online-Workshopkonzept haben wir bereits bei einem erfolgreich durchgeführten Ferienprogramm in den Osterferien mit Schüler/innen gesammelt. Die positiven Rückmeldungen seitens der Lernenden bestärken uns darin, unser Angebot nicht nur für ortsnahe Schulen anzubieten, sondern Schulen in ganz Deutschland den Zugang zu unserem Workshopangebot zu ermöglichen. In den Sommerferien wird es weitere Ferienangebote<sup>5</sup> unsererseits geben.

Wir freuen uns auf viele Anfragen!

<sup>1</sup> CAMMP steht für das Projekt Computational and Mathematical Modeling Program

<sup>2</sup> Jupyter Notebook ist eine Open-Source-Webanwendung, mit der man Dokumente erstellen und austauschen kann, die Live-Code, Gleichungen, Visualisierungen und beschreibenden Text enthalten (Quelle: [jupyter.org](https://jupyter.org)).

<sup>3</sup> Siehe auch [discord.com](https://discord.com)

<sup>4</sup> Siehe auch [mattermost.com](https://mattermost.com)

<sup>5</sup> Termine und weitere Informationen dazu unter: [www.cammp.online/186.php](https://www.cammp.online/186.php)

### CAMMP goes online

In order to reduce the risk of infection with the corona virus, the Ministry of Education, Youth and Sports cancelled all extracurricular events until the summer holidays. This cancellation/ban includes the CAMMP days of the project "Computational and Mathematical Modeling Program" at KIT. Thus, the CAMMP team was challenged to establish and implement online courses.

## +++Neues aus den SCC-Abteilungen+++

### Neue Leitung der Abteilung Netze und Telekommunikation



Seit 1. Juli 2020 ist Klara Mall Leiterin der Abteilung Netze und Telekommunikation (NET). Frau Mall übernahm im November 2017 die Stellvertretung des langjährigen Abteilungsleiters Reinhard Strebler, der Ende 2018 seinen Ruhestand antrat (SCC-News 2/2018). Seit Reinhard Strebler's Weggang wurde sie zunächst mit der technisch-fachlichen Leitung der Abteilung betraut. So konnte sich Frau Mall für eine Übergangsphase mit den Kolleginnen und Kollegen auf die anstehende Realisierung von zentral bedeutenden Projekten konzentrieren. Klaus Scheibenberger, Leiter der Abteilung Systeme und Server (SYS), hatte während dieser Zeit die administrative Leitung von NET inne.

Da die Umsetzungsprojekte von NET meist mit SYS und umgekehrt intensiv abgestimmt werden müssen, vertreten sich Klara Mall und Klaus Scheibenberger seit 1. Juli 2020 gegenseitig.

Klara Mall arbeitete schon zwischen 2003 und 2005, während ihres Mathematikstudiums an der Universität Karlsruhe, als Studentische Hilfskraft im dortigen Rechenzentrum und konnte so die Netzwerke der Vorgängereinrichtungen des KIT in der Praxis kennenlernen. Seit Juli 2008 ist sie Mitarbeiterin der Abteilung Netze und Telekommunikation am SCC. Seit den Anfängen des KIT war sie am Aufbau und Betrieb der notwendigen Netzwerktechnologien für eine der größten Wissenschaftseinrichtungen Europas maßgeblich beteiligt. Klara Mall ist mit ihrer Expertise auch in unterschiedlichen Projekten national und international engagiert.

Das SCC informiert seit 1. Juli 2020 auch über einen eigenen Twitterkanal (@SCC\_KIT) – unter anderem zu Wissenschaftlichem Rechnen, Forschungsdatenmanagement, IT-Diensten, Events und allgemeinen Neuigkeiten.

## IMPRESSUM

**SCC news**  
Magazin des Steinbuch Centre  
for Computing

**Herausgeber**  
Präsident Professor Dr.-Ing.  
Holger Hanselka  
Karlsruher Institut für Technologie (KIT)  
Kaiserstraße 12  
76131 Karlsruhe

**Anschrift**  
Steinbuch Centre for Computing (SCC)  
Karlsruher Institut für Technologie (KIT)  
Redaktion SCC-News  
Zirkel 2  
76131 Karlsruhe  
oder:  
Hermann-von-Helmholtz-Platz 1  
76344 Eggenstein-Leopoldshafen  
Fax: +49 721 608-24972

**Redaktion**  
Achim Grindler (verantwortlich),  
Karin Schäufele, Andreas Ley  
E-Mail: [redaktion@scc.kit.edu](mailto:redaktion@scc.kit.edu)

**Gestaltung, Satz und Layout**  
Heike Gerstner, Hella Grolmus  
AServ – Crossmedia – Grafik  
Karlsruher Institut für Technologie (KIT)  
Hermann-von-Helmholtz-Platz 1  
76344 Eggenstein-Leopoldshafen

**Titelfoto**  
Amadeus Bramsiepe

**Fotos**  
Achim Grindler, Simon Raffener,  
Maren Hattebuhr, Ulrich Weiß

**Druck**  
Systemedia GmbH, 75449 Wurmberg

**Erscheinungstermin dieser Ausgabe**  
August 2020

[www.scc.kit.edu/publikationen/scc-news](http://www.scc.kit.edu/publikationen/scc-news)

**Der Nachdruck und die elektronische Weiterverwendung sowie die Weitergabe von Texten und Bildern, auch von Teilen, sind nur mit ausdrücklicher Genehmigung der Redaktion gestattet.**



Karlsruher Institut für Technologie (KIT)  
Steinbuch Centre for Computing (SCC)

ISSN: 1866-4954

[www.scc.kit.edu](http://www.scc.kit.edu)  
[www.scc.kit.edu/twitter](http://www.scc.kit.edu/twitter)  
[contact@scc.kit.edu](mailto:contact@scc.kit.edu)